UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA

Influence of histone octamer core distortion on nucleosome thermal mobility

Alexandre Dias Cirilo

Mestrado em Bioquímica
Especialização em Bioquímica

Dissertação orientada por:
Dra. Silvija Bilokapić-Halić
Prof. Dra. Margarida Gama-Carvalho

2018

Acknowledgements

I would first like to thank my external supervisor Dr. Silvija Bilokapić-Halić from the Ludwig-Maximilians-Universität München (LMU) in Munich, Germany. I learned a lot throughout this academic year and am sure that having been driven to think by myself, perform and critically analyse the experiments was the best way to grow as a scientist. I would also like to thank Prof. Dr. Mario Halić for accepting me into the Halić lab at the Gene Center of the LMU and allowing me to do my master thesis project for a whole academic year. Moreover, I would like to thank everybody that I worked with during my stay, especially the members of the lab. Thank you for helping me whenever I required assistance and for all the support.

I would also like to acknowledge Prof. Dr. Margarida Gama-Carvalho of the Faculty of Sciences at the University of Lisbon who gladly accepted to be my internal supervisor and assisted me whenever needed.

Furthermore, I want to reference the Erasmus+ programme ant its respective grant allocation without which I would not have been able to pursue my thesis project abroad.

Finally, I am deeply grateful to my parents for allowing me to pursue my objectives and ambitions, for always being present and for all the support and endless love. This work would not have been possible without them.

Abstract

The nucleosome is the fundamental repeating subunit of eukaryotic chromatin. The histones, hallmarks of the nucleosome, are small basic proteins subdivided into two super families: linker (H1) and core histones (H2A, H2B, H3 and H4). Core histones assemble together to form the histone octamer which then, together with double-stranded DNA, embodies the complex structure of the nucleosome. Most importantly, the nucleosome allows for the tight packaging of the large genome into the relatively constricted space available inside the nucleus by folding into a series of higher-order molecular structures. Furthermore, it allows for the fine-tuned temporal and spatial regulation of several DNA-templated processes, such as DNA transcription, replication and repair, unequivocally required for the cells to thrive and exert their functions. Since their discovery in the 1960s, histone post-translational modifications (e.g. methylation, phosphorylation and acetylation) have been shown to affect gene expression regulation, with some modifications being associated with gene activation, whilst others have been associated with gene silencing. Interestingly, several studies have demonstrated the existence of a "histone code" which states that the presence of post-translational modifications on histones affect binding affinities of chromatin-binding proteins leading to altered transcriptional states of the underlying DNA. The epigenetic information encoded via these modifications occurring on histones can be passed on to the next generation of cells, acting as an additional layer of information that can be stored inside the cell, further increasing the complexity and versatility of the genetic material.

It has been shown that the nucleosome exhibits shifting when incubated at high temperatures towards the DNA ends in regard to the nucleosomal DNA. In this context, the current project aims to observe the influence of histone octamer core distortion on nucleosome thermal mobility. To do so, several histone mutants were prepared, albeit only one mutant nucleosome was assembled, apart from the wild-type nucleosome. This work corroborates the observed thermally-driven shifting pattern of wild-type nucleosomes. The H4_V43C H3_F104C mutant nucleosome assembled in this project has shown itself to increase nucleosome structural integrity, possibly through the establishment of a disulphide bridge between L1 (loop 1) of core histone H4 and α2 (α-helix 2) of core histone H3, conversely, hindering nucleosome mobility. This suggests that for nucleosome to slide along DNA in a noncatalyzed fashion, there must be a requirement of histone octamer plasticity.

Moreover, structural analysis of Fkbp39, a putative histone chaperone in *Saccharomyces pombe*, both alone and in complex with (H2A-H2B) and with $(H3\text{-}H4)_2$ was attempted with electron cryomicroscopy, though unsuccessful. Further optimization strategies are likely required to stabilize the putative complexes.

Lastly, several chromatin-binding proteins were screened for binding to the nucleosome, as a complementary experiment. Hpf1, Parp2 and Alf from *Homo sapiens* seem like the best candidates for more in-depth studies to fully disclose their binding affinities to chromatin.

Keywords: Nucleosome, Histone code, PTMs, Epigenetics, Chromatin

Resumo

O nucleossoma é a subunidade fundamental da cromatina eucariótica. As histonas, proteínas características do nucleossoma, são pequenas proteínas básicas subdivididas em duas superfamílias: *linker* (H1) e *core* (H2A, H2B, H3 e H4). As histonas *core* oligomerizam para formar o octamero de histonas que, conjuntamente com ADN de cadeia dupla, estabelecem a estrutura complexa do nucleossoma. A estrutura do nucleossoma foi inicialmente identificada em 1975 recorrendo à digestão da cromatina por uma nuclease micrococal, sendo que apenas em 1997 foi divulgada a estrutura do nucleossoma obtida por difração de raios-X com uma resolução de 2.8 Å, e mais tarde, em 2002, com uma resolução de 1.9 Å, o que revela ser mais do que suficiente para as cadeias laterais e a cadeia principal serem observadas com elevado grau de confiança. As histonas partilham um domínio entre elas: o *histone fold*, constituído por cerca de 70 resíduos de aminoácidos, capaz de facilitar a heterodimerização das histonas e que se encontra localizado na região C-terminal e consiste em três hélices-α (α1, α2 e α3) ligadas entre si por dois *loops* (L1 e L2) flanqueando a hélice-α 2. As histonas são constituídas por um domínio globular e uma região N-terminal flexível rica em resíduos de lisina e arginina. Estas proteínas heterodimerizam com o auxílio de interações hidrofóbicas, formando o *handshake motif*. No contexto celular, o nucleossoma permite o empacotamento do genoma no espaço relativamente limitado disponível dentro do núcleo, através do *folding* numa série de estruturas moleculares de ordem gradualmente superior. Mais ainda, permite a regulação temporal e espacial de vários processos modelados pelo ADN, como por exemplo a transcrição, a replicação e o reparo de ADN, necessários para que as células prosperem e exerçam as suas funções.

Desde a sua descoberta na década de 1960, as modificações pós-traducionais de histonas (*e.g.* metilação, fosforilação e acetilação) foram associadas à regulação da expressão génica, com algumas modificações sendo associadas à ativação de genes, enquanto outras foram associadas ao seu silenciamento. Vários estudos demonstraram a existência de um "código de histonas" que estipula que diferentes modificações pós-traducionais nas histonas afetam as afinidades de ligação de proteínas capazes de estabelecerem uma ligação à cromatina, levando a estados transcricionais alterados do ADN subjacente. A informação epigenética codificada por estas modificações que ocorrem nas histonas pode ser passada para a próxima geração de células, agindo como uma camada adicional de informação que pode ser armazenada dentro da célula, aumentando ainda mais a complexidade e a variabilidade do material genético.

Foi previamente demonstrado que o nucleossoma exibe um deslocamento, quando incubado a altas temperaturas, para os terminais da molécula de ADN que está enrolada em volta do nucleossoma. Neste contexto, o presente projeto visa observar a influência da distorção do octamero de histonas na mobilidade térmica do nucleossoma. Deste modo, vários mutantes de histonas foram preparados, embora apenas um nucleossoma mutante tenha sido gerado, além do nucleossoma *wild-type*. Para a obtenção de histonas recombinantes com elevado grau de pureza, recorreu-se a técnicas cromatográficas. Em primeiro lugar, foram expressas histonas recombinantes (*wild-type* e mutantes) em células *Escherichia coli* BL21 (DE3) Rosetta competentes a partir de plasmídeos *in house*. As histonas *core* H2A e H2B, bem como H3 e H4, foram coexpressas, uma vez que a coexpressão destes dois pares de proteínas reduz o tempo necessário para a sua síntese, para além de produzir complexos solúveis passíveis de serem purificados por sucessivas técnicas experimentais. Posteriormente, executou-se a purificação das mesmas através do uso de uma cromatografia de afinidade seguida duma cromatografia de troca iónica.

Após todas as histonas terem sido geradas em grandes quantidades e elevado grau de pureza, efetuou-se a montagem do octamero. Paralelamente, foi sintetizada a sequência Widom 601, previamente descrita na literatura como sendo uma sequência de muito alta afinidade para o nucleossoma, via PCR usando primers apropriados. Neste trabalho foram geradas duas variantes da sequência supramencionada: curta (147 pb) e longa (227 pb). Uma vez obtidos todos os componentes necessários para a montagem do nucleossoma, recorreu-se à reconstituição do mesmo através de diálises (com gradiente de salinidade). Por fim, os nucleossomas gerados foram usados para serem efetuados *thermal shift assays* e *native gel shift assays*. Os *thermal shift assays* foram usados para se observar o deslocamento do nucleossoma ao longo do ADN (baseado na sequência Widom 601) e o efeito da distorção do octamero na mobilidade do nucleossoma induzida pela temperatura. Quanto aos *native gel shift assays*, um ensaio experimental rápido e sensível para a deteção de interações proteína-ácido nucleico, foram usados para observar a afinidade de ligação de diversas proteínas com ligação putativa à cromatina.

Este trabalho corroborou o padrão observado de deslocamento de nucleossomas *wild-type* para as extremidades do ADN associado ao mesmo. O nucleossoma mutante H4_V43C H3_F104C concebido neste projeto demonstrou estar associado ao aumento da integridade estrutural e rigidez do nucleossoma, possivelmente através do estabelecimento de uma ponte dissulfeto entre L1 (*loop* 1) da histona H4 e α2 (hélice-α 2) da histona H3. O estudo do impacto da plasticidade do octamero de histonas, consequência das diversas modificações que podem ocorrer nos seus resíduos de aminoácidos e que levam à distorção do *core* de histonas, é relevante no sentido de tentar entender o seu papel na regulação da expressão génica. Uma dada mutação irá, hipoteticamente, reforçar a estrutura do nucleossoma, potencialmente impossibilitando o acesso do complexo basal de transcrição, devido à alteração da capacidade de mobilidade do nucleossoma em relação ao ADN, e silenciando o(s) gene(s) codificados no ADN subjacente que se encontra "blindado" pelo nucleossoma, e vice-versa. Também será possível que alguns complexos capazes de remodelar a cromatina possam tirar partido da alteração da integridade estrutural conferida pelas diversas mutações ou pela presença de variantes de histonas. Numa visão geral, os resultados obtidos neste projeto demonstraram que é necessária uma certa plasticidade do octamero de histonas para se observar o movimento não-catalisado do nucleossoma.

Adicionalmente, foi tentada a análise da Fkbp39, uma peptidilprolil isomerase com um domínio NPL (*nucleoplasmin-like*) conservado na região N-terminal, que se pensa ser uma chaperona de histonas em *S. pombe* e que é responsável pelo aumento da taxa de isomerização cis-trans das prolinas na cauda N-terminal (região N-terminal flexível) da histona H3. A análise foi feita por criomicroscopia eletrónica, tanto com a Fkbp39 não-complexada, como em complexo com (H2A-H2B) e com $(H3\text{-}H4)_2$. Nesse sentido, expressou-se a proteína Fkbp39 em células *E. coli* BL21 (DE3) Rosetta competentes e foi executada a sua purificação com elevado grau de pureza através do uso de diferentes técnicas cromatográficas (cromatografias de afinidade, de troca iónica, e de exclusão molecular). Embora as purificações tenham sido bem-sucedidas, nem a proteína Fkbp39 isolada, nem em complexo com o dímero (H2A-H2B) ou com o tetramero $(H3\text{-}H4)_2$ foram passíveis de serem observadas por criomicroscopia electrónica. É possível que, devido à presença de uma região central dinâmica intrinsecamente desordenada, estratégias sucessivas de otimização sejam necessárias para estabilizar os complexos putativos.

Por último, várias proteínas putativas de ligação à cromatina foram rastreadas quanto à ligação aos nucleossomas sintetizados neste projeto recorrendo a *native gel shift assays*. Hpf1 (*Histone PARylation factor 1*), Parp2 (*Poly [ADP-ribose] polymerase 2*) e Alf (*TFIIA-alpha and beta-like factor*), provenientes de *Homo sapiens*, demonstraram ser os melhores candidatos para estudos estruturais posteriores mais aprofundados para averiguar com maior confiança as suas afinidades de ligação à cromatina.

Palavras-chave: Nucleossoma, Código de histonas, Modificaçóes pós-traducionais, Epigenética, Cromatina

Index

List of Abbreviations

Table 0.1 List of abbreviations

Abbreviation	Designation
AC	Affinity chromatography
ACF	ATP-utilizing chromatin assembly and remodelling factor
Amp	Ampicillin
Cam	Chloramphenicol
ChIP	Chromatin immunoprecipitation
c-NCP	Centred nucleosome core particle
Cryo-EM	Electron cryomicroscopy
D. melanogaster	*Drosophila melanogaster*
DNA (ADN)	Deoxyribonucleic acid (Ácido desoxirribonucleico)
dNTP	Deoxyribonucleotide triphosphate
DSB	Double-strand break
dsDNA	Double-stranded deoxyribonucleic acid
E. coli	*Escherichia coli*
EM	Electron microscopy
ep-NCP	End-positioned nucleosome core particle
FKBD	FK506-binding domain
H. sapiens	*Homo sapiens*
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
His	Histidine
IB	Inclusion body
IDP	Intrinsically disordered proteins
IEX	Ion-exchange chromatography
Kan	Kanamycin
LB	Lysogeny broth
MACS	Multiangle light scattering
MMTV	Mouse mammary tumour virus
N/A	Not available
NCP	Nucleosome core particle
NMR	Nuclear magnetic resonance
NPL	Nucleoplasmin-like
OD_{600}	Optical density at a wavelength of 600 nm
P. falciparum	*Plasmodium falciparum*
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PPIase	Petptidyl-prolyl cis-trans isomerase
PTM	Post-translational modification
ref(s).	Reference(s)
RSF	Remodelling and spacing factor
RT	Room temperature

S. pombe	*Schizosaccharomyces pombe*
SANS	Small-angle neutron scatterin
SAXS	Small-angle X-ray scattering
SE AUC	Analytical ultracentrifugation: sedimentation equilibrium
SUMO	Small ubiquitin-related modifier
t	Time
T_m	Melting temperature
WT	Wild-type
X. laevis	*Xenopus laevis*

List of Tables

List of Figures

1. Introduction

1.1. Historical contextualization and overview of histones

Nearly half a century ago, the foundations for present-day chromatin research were erected when Kornberg and colleagues proposed that “chromatin structure is based on repeating units of eight histone molecules and about 200 DNA base pairs”[1]. Although the existence of histones, as a biological entity, was revealed in the late 19th century (1884), it wasn’t until 1974, with the advent of modern biochemistry that brought advanced extraction and purification methods, that the actual stoichiometry and structural properties of the histone octamer were made public to the scientific community, ultimately leading to the uncovering of an $(H3\text{-}H4)_2$ tetramer and an (H2A-H2B) dimer[1]. Later on, the structural and spatial arrangements of these small basic proteins were elucidated through crystallographic analyses of the octamer, showing that the core histones, albeit sharing low sequence identity, revealed a characteristic structural motif termed the "histone fold" (approximately 70 amino acid residues) – a domain capable of facilitating heterodimerization found near the C-terminus of core histones and consisting of three α-helices (α1, a longer middle α2, and α3) connected by two loops (L1 and L2) in between them like so: α1-L1-α2-L2-α3 (see Figure 1.1)[2]. Dimerization of histones occurs in an anti-parallel orientation, and it has been shown that the histone core binds the nucleosomal DNA at 14 superhelix locations, making up more than 120 direct atomic interactions[3]. The histone octamer is formed by a central $(H3\text{-}H4)_2$ tetramer flanked by two (H2A-H2B) dimers. These four dimers are arranged in a two-fold symmetry axis, termed the dyad, which intersects with the middle of the DNA fragment. These histones, highly conserved proteins among eukaryotes (see Figure 9.1), consist of a globular domain and a flexible and charged N-terminal (also termed the "histone tail"; see Figure 1.2) rich in charged lysine and arginine residues, they assemble into heterodimers via hydrophobic interactions by forming a "handshake motif" (see below in Figure 1.1) and folding the loops into β-bridges[2].

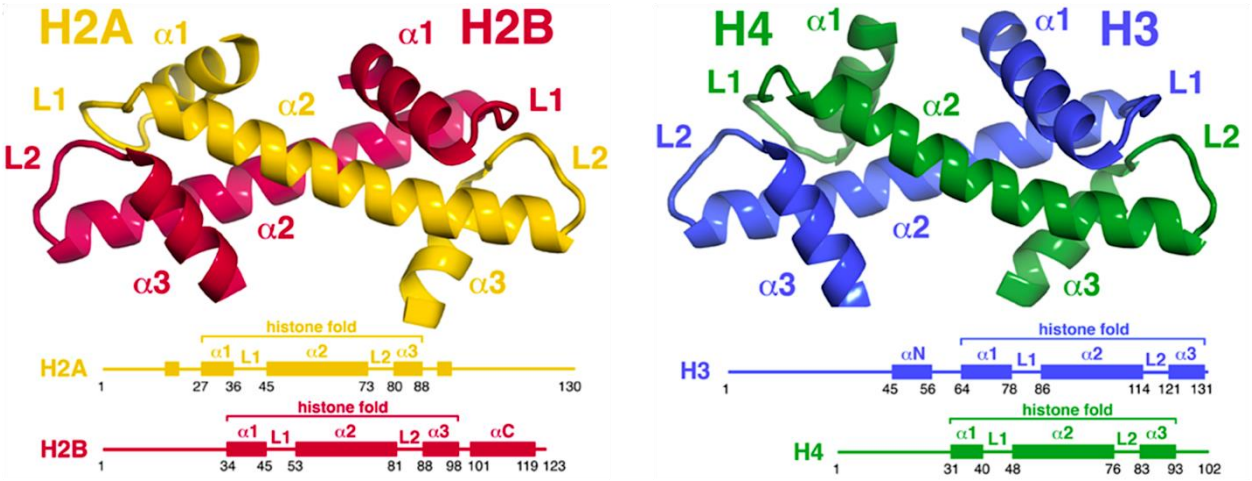


Figure 1.1 The histone fold. Secondary structures are depicted and labelled accordingly. Linear schemes illustrate the presence of the histone fold in all four core histones. Adapted from ref. 4.

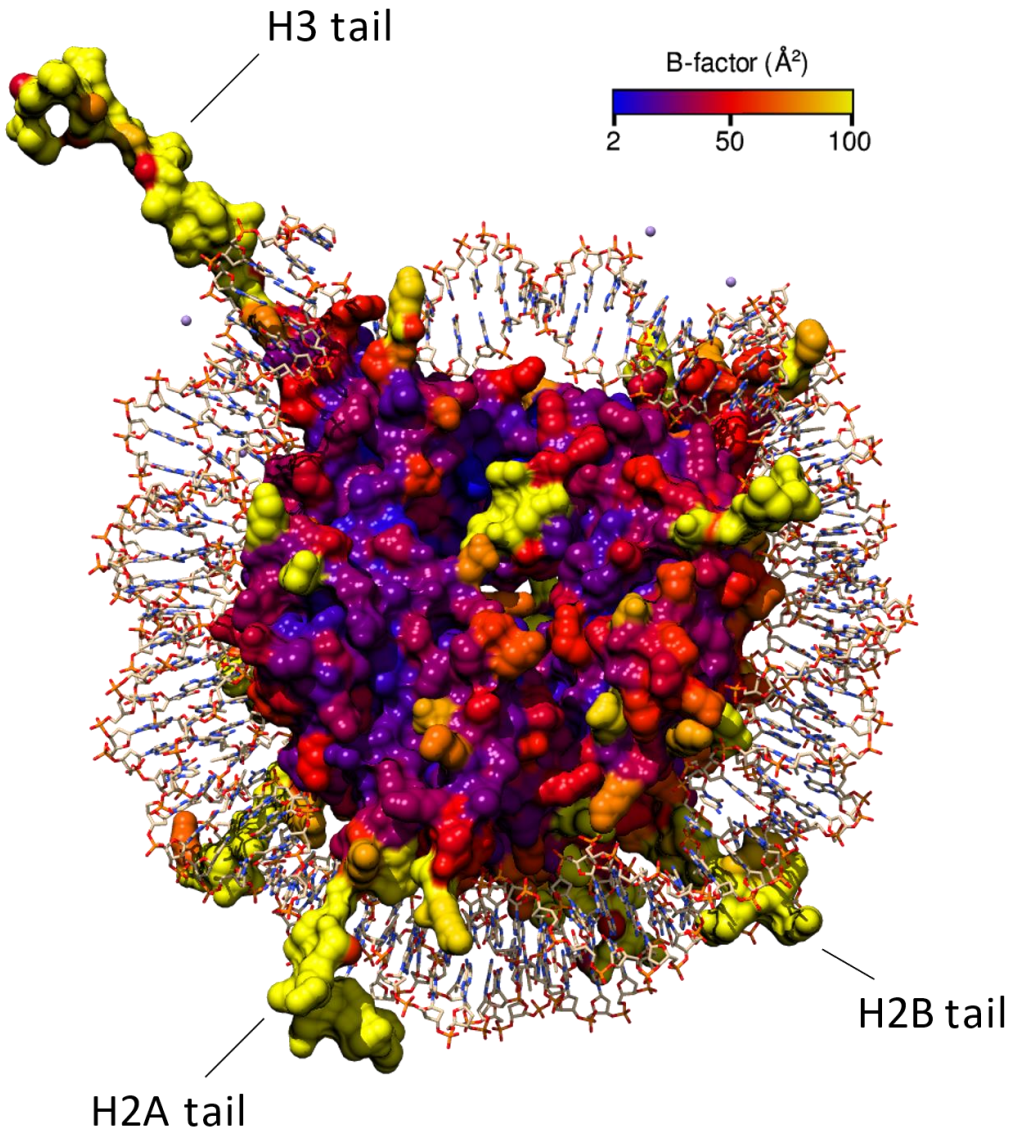


Figure 1.2 Graphical representation of the nucleosome with its core histones coloured by B-factor. From the lowest to the highest order b-factor: blue, red, and yellow. The B-factor (also termed "temperature factor" or "Debye-Waller factor) is a good indicator of the electron density spreading extent, thus being useful for determining disorder in particular regions of the molecule (generally above 60 Å^2), as well as being useful for the identification of potential errors in model building. Visible histone tails are annotated accordingly. Representation generated using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081; PDB: 1AOI)[84].

1.2. The nucleosome – structural properties and role in genetic material packaging

Inevitably, when talking about histones, the term “nucleosome” cannot be overlooked. By definition, the nucleosome refers to the fundamental repeating subunit of eukaryotic chromatin, which is nowadays considered to be dynamic and polymeric in nature. The nucleosome (see Figure 1.3) consists of two copies of each core histone, i.e. histone H2A, histone H2B, histone H3, and histone H4, which form the histone octamer, and has a diameter of approximately 11 nm. DNA is wrapped around the latter in roughly two turns, forming what came to be known as the "nucleosome core particle" (NCP) with a molecular weight of around 200 kDa and a diameter of approximately 100 Å. Each NCP is linked to the adjacent NCP through a linker DNA – ds DNA that holds the two NCPs together and is susceptible to endonuclease degradation – with a length ranging from 10 to 90 base pairs resulting in the well-known chromatin polymer. Additionally, a small portion of this linker DNA (approximately 20 base pairs) is typically bound to the linker histone H1 (which has many reported variants), forming the chromatosome, i.e. the NCP with the linker histone. Adding the remaining linker DNA to the chromatosome leads to the formation of the nucleosome. From data collected around the assembly and disassembly of nucleosomes, several intermediate structures with suboctameric stoichiometries can be inferred: the hexasome, lacking one (H2A-H2B) dimer, and the tetrasome, lacking two (H2A-H2B) dimers. Structural analyses have confirmed that the hexasome shows standard nucleosome architecture, albeit harbouring only a 110 base pairs DNA. In the wake of the discovery of these non-canonical nucleosome states, an additional complex has been suggested: the hemisome, which consists of one single copy of each core histone and incorporates other proteins and can accommodate reverse DNA supercoils. Besides these complexes, several additional variants have been proposed later on, highlighting the idea that the nucleosome exhibits a dynamic and polymorphic behaviour in vivo[4,5].

The structure of the nucleosome was solved swiftly after Kornberg’s proposal, in 1975, resorting to the micrococcal nuclease digestion of chromatin which cleaved the linker DNA, leaving each single nucleosome unit with DNA of just about 146 base pairs coiled around it. The ensuing structure, the nucleosome core particle, was found to be remarkably conserved and could be subjected to crystallization and subsequent structural biology methods for structural elucidation[1]. Although higher resolution structures of the nucleosome core particle were only obtained some years later, with a 7 Å X-ray structure being solved in 1984[6], a near-atomic resolution of 2.8 Å X-ray structure (see Figure 1.4) solved one decade later, in 1997[7], and a 1.9 Å X-ray structure published in 2002[8].

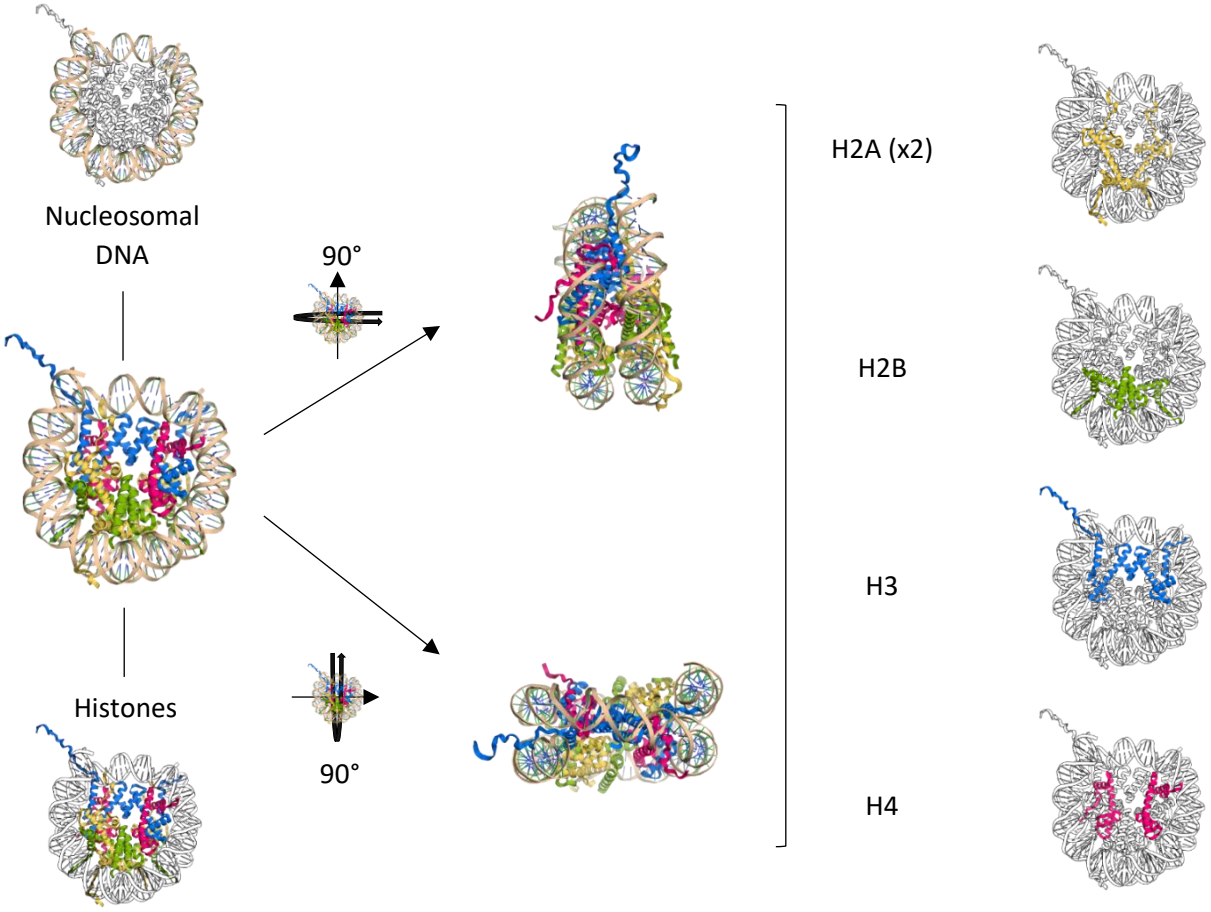


Figure 1.3 Overview of the nucleosome and its components in *X. laevis*. Each core histone is depicted accordingly to one colour (H2A: yellow, H2B: green, H3: blue, H4: magenta). Representations generated using PyMOL (PDB: 1AOI; The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

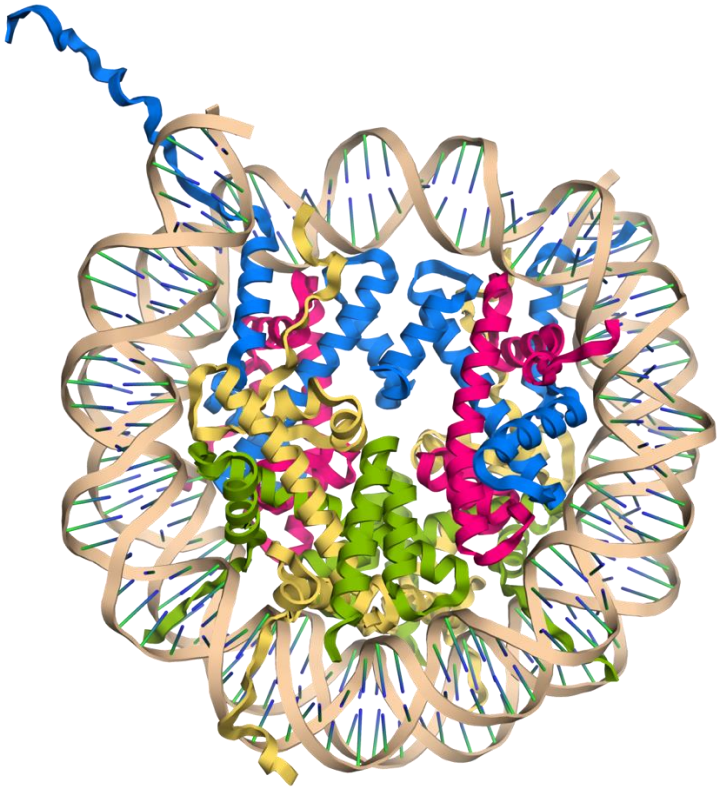


Figure 1.4 2.8 Å solved structure of the nucleosome core particle. Core histones H2A in yellow, H2B in green, H3 in blue, and H4 in magenta. Nucleosomal DNA coloured in beige (helices only). Representation generated using PyMOL (PDB: 1AOI; The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

The relevance of the nucleosome is justified by its swift and efficient control over the structure and activity of chromatin, which is the primary agent for the plethora of cellular processes involving DNA that occur in the nucleus[9]. In addition to the requirement of a tight packaging of the genetic material inside the nucleus, there is the need for an easily accessible genome by factors that perform the various DNA-templated processes (e.g. transcription, recombination and repair). This task is attributed to the histones which bind with DNA very tightly due to the strong electrostatic interactions between the intrinsic positive charge of histones and the negatively charged DNA. In regard to the accessibility of DNA within chromatin, multiple entities must work together in order to "mould" chromatin, notably, histone-modifying enzymes (e.g. histones acetyltransferases and histone methyltransferases), transcription factors and chromatin remodelling complex (e.g. SWI/SNF). In a broader view, one can state two main mechanisms by which DNA can be accessed: 1) through the enzymatic modifications of histones, e.g. methylation and acetylation; 2) by the action of chromatin remodelling complexes capable of exposing DNA sequences.

In the greater scheme of things, nucleosomes fold up to form the 30 nm diameter chromatin fibre, commonly known as the "beads on a string" structure, which rearranges itself in higher-order structured loops of around 300 nm in length. The latter is then compressed to form a 250 nm wide fibre, which is then coiled into the chromatid of a chromosome. This packaging mechanism allows for an optimal compaction and decrease in occupied space, which is really useful considering the relatively small unoccupied space available inside the nucleus[10].

1.2.1.Core histones

The histone octamer core has been shown to exist as an intact biological entity at high ionic strengths, even without nucleosomal DNA or any crosslinking agents. The latter is not observed at physiological ionic strengths, instead, the histone octamer dissociates into the $(H3\text{-}H4)_2$ tetramer and two (H2A-H2B) dimers. As said beforehand, the four core histones (H2A, H2B, H3, and H4) share a common and highly conserved structural motif: "the histone fold". Several additional proteins have been observed to contain this motif, suggesting it may be hiding other roles. H2A and H3 contain additional helices at the N-terminus, and H2B contains an additional helix at the C-terminus (see Figure 1.1). Additionally, core histones also have flexible N-terminal tails that contain a basic region. It is thought that these tails are disordered in the absence of DNA, since they are not visible in the octamer structure[11].

1.2.2.Linker histone H1/H5

The linker histone H1 (H5 is a variant of linker histone H1 in which many lysine residues have been replaced by arginine residues) binds to the nucleosome and promotes its packaging in a higher-order structure. Experiments using micrococcal nuclease to digest chromatin have shown that the latter originated the chromatosome, a metastable intermediate consisting of just about 165 bp DNA coiled around the octamer, and the linker histone H1. Further digestion by the same agent resulted in the dissociation of linker histone H1 and the digestion of an additional 20 bp of DNA, originating, as said beforehand, the nucleosome core particle (with ~146 bp of DNA). When there was no H1, the chromatosome could not be observed. Restoring H1 reintroduced the chromatosome. Hence, this experiment suggested that the linker histone H1, in addition to its binding to the nucleosome, could also protect an additional 20 bp nucleosomal DNA. Moreover, it has also been observed that the tails of the linker histone span over half the molecule[11].

1.2.3.Histone variants

Histone variants have been observed for all histones, except for core histone H4. They usually differ by a few amino acid residues or by the inclusion of specific domains. The presence of histone variants introduces flexibility and versatility to chromatin, providing additional means to regulate the plethora of DNA-templated processes. These variants are classified into two main categories: replicative (or canonical) histones and replacement histones. And they can affect chromatin organization, histones´ PTMs, as well as the binding to specific partners through the modification of their biochemical properties. Replicative histones, encoded by multiple gene copies organized in tandem and devoid of introns, show a peak of expression in the S phase of the cell cycle, and their incorporation into chromatin is coupled to DNA synthesis. On the other hand, replacement histones are mainly expressed throughout the cell cycle, sometimes in a tissue-specific manner, and encoded by single genes. Their incorporation into chromatin can occur throughout the cell cycle and is independent of DNA synthesis. They are also particularly enriched in highly transcribed regions, centromeres and telomeres[4].

1.2.4.Nucleosome assembly

As said earlier, the nucleosome is maintained by the strong attraction between positively charged histones (due to being enriched in lysine and arginine residues) and the negatively charged DNA (due to the phosphate groups). In vitro, mixing DNA and histones leads to the formation of insoluble aggregates instead of nucleosomes, suggesting the presence of additional factors that mediate the interactions between DNA and histones. Some pioneering studies from the late 1970s presented proteins that acted as chaperones and were responsible for the prevention of aggregate formation at physiological salt concentrations, as well as having a role in nucleosome assembly. In vitro, these histone chaperones can facilitate nucleosome assembly by preventing the formation of aggregates between histones and DNA. In vivo, they bind histones, prevent non-specific interactions, and promote interactions with

specific components such as DNA and other proteins. Today, the consensus asserts that all free histones are bound to histone chaperones in the cellular context. Generally, histone chaperones exhibit multiple functions, namely: 1) transport of newly synthesized histones from the cytoplasm to the nucleus; 2) transport histones to histone-modifying proteins to introduce PTMs; 3) storage of free histones; 4) deposition of histones on DNA; 5) removal of histones from the nucleosome and/or DNA. These molecular chaperones specific for histones collaborate with chromatin remodelers to achieve an efficient and fine-tuned assembly and disassembly of chromatin. In vitro, free (H2A-H2B) and (H3-H4) are observed as heterodimers at physiological conditions (ionic strength and pH), even though some $(H3\text{-}H4)_2$ tetramers can form. The currently accepted model for nucleosome assembly states that, firstly, either two dimers or one tetramer of $(H3\text{-}H4)_{(2)}$ is deposited onto DNA, forming the tetrasome. As mentioned earlier, the tetrasome consists of DNA (almost one turn) coiled around one $(H3\text{-}H4)_2$ tetramer. Afterwards, two (H2A-H2B) dimers are added to the structure, allowing for the remaining DNA to coil around the newly formed histone octamer. Nucleosome disassembly is thought to be the complete reverse process of nucleosome assembly. This stepwise model is thought to be executed by histone chaperones[4].

1.2.5. Histone chaperones

Exhibiting multiple functions, histones chaperones, as suggested by their name, aid histones in various important processes, such as their transport from the cytoplasm to the nucleus, their storage in the cellular context, their presentation to histone-modifying enzymes for the introduction of PTMs, and naturally, their assembly/disassembly. Some histone chaperones are known to be histone-specific, e.g. chaperones specific for (H2A-H2B) dimers. Additionally, histone chaperones can act at specific cellular events: during DNA replication or when DNA is not being synthesized, i.e. during transcription or not, being termed accordingly as "replication dependent" and "replication independent", respectively[4]. They can bind histones for a limited time to be modified a posteriori by histone modifying enzymes, or alternatively, keep them in an inactive state (e.g. to avoid binding of misfolded non-canonical histones to DNA, a function observed in nucleoplasmin)[12].

The first histone chaperone was identified in 1978, purified from *X. laevis*, it was termed "nucleoplasmin"[13], being extensively characterized over the years as a (H2A-H2B) histone chaperone. In its particular case, binding to positively charged histones seems to be favoured by the protein's abundance of phosphorylated residues and to the presence of acidic loops[12]. Nowadays, several histone chaperones have been reported. Namely, for example, the chromatin assembly factor 1 (CAF-1) – a replication-dependent deposition factor that promotes the assembly – and the histone regulator A (HIRA) – a protein which is thought to execute the same function as CAF-1, albeit in a replication-independent fashion. Histone chaperones are thought to have two prominent facets: 1) storage and specificity; 2) deposition and coupling with DNA-templated processes. They are responsible for escorting histones from their synthesis to their death, being involved in their provision, localization, eviction and degradation, and can operate through an "escort network", exhibiting characteristic structural conformations that allow their recruitment to specific genomic locations and linking to particular cellular mechanisms. Noteworthily, there seems to be a great interest in studying the role of histone chaperones in development and disease[13].

Naturally, before the assembly of histones occurs, their transport from the cytoplasm to the nucleus must be assured. Regarding (H2A-H2B), this task is most probably performed by Nap1 (Nucleosome assembly protein 1), a cytosolic histone chaperone associated with the shuttling between the nucleus and the cytoplasm. On the other hand, the transport of (H3-H4) has only been elucidated more recently. It is likely that the aforementioned shuttling involves a complex framework of interactions coordinated by multiple histone chaperones in a sequential workflow[4].

1.2.5.1. FK506-binding protein 39 kDa

As a side project, we wanted to assess if Fkbp39 (FK506-binding protein 39 kDa), also termed "Ani1"[14,15] in *S. pombe* (commonly known as “fission yeast”), could form a complex in vitro with heterologously-expressed WT (H2A-H2B) and $(H3\text{-}H4)_2$ from *X. laevis*.

Fkbp39, a PPIase localized in the nucleolus[16] responsible for the increase of the rate of cis-trans isomerization at prolines on the histone H3 N-terminal tail, consists of a NPL core domain conserved in eukaryotes and fungi located on the N-terminal side of the protein, and a FKBD domain located on the C-terminal linked to the latter via central dynamic highly charged intrinsically disordered regions (formed by acidic and basic stretches localized between the two aforementioned domains, NPL and FKBD)[17].

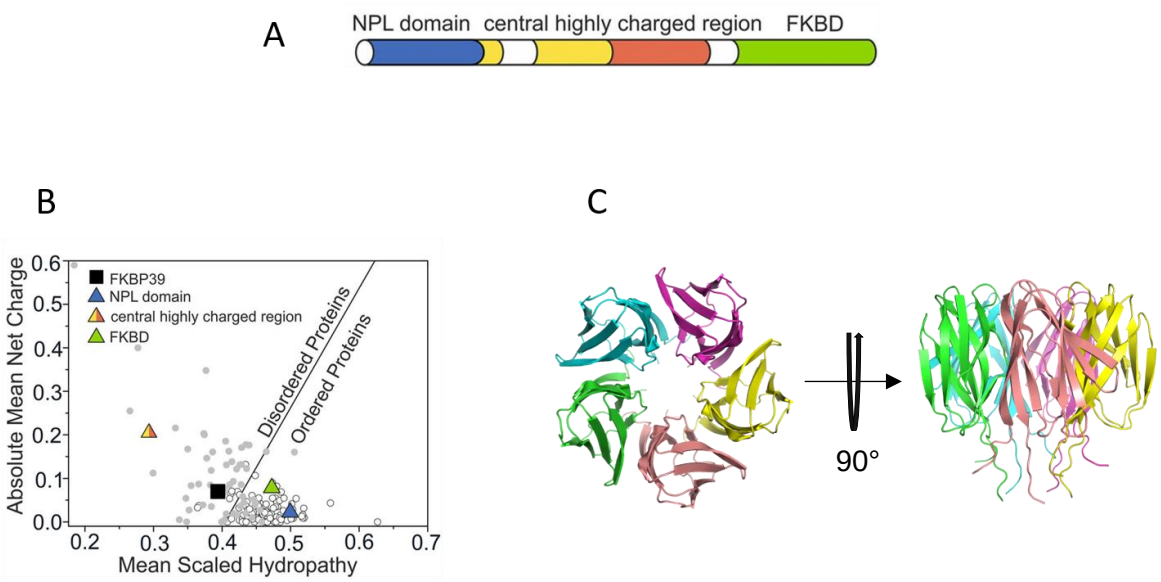


Figure 1.5 Fkbp39 structural organization. (A) Structural organization of FKBP39’s domains obtained by Pfam67 (ref. [18]). N-terminal NPL domain coloured in blue; central highly charged region coloured in yellow for acidic stretches and in red for basic stretches, and the C-terminal FKBD is coloured in green. (B) Uversky plot of Fkbp39. Globular proteins represented by white circles and disordered proteins represented by grey circles. Full-length Fkbp39 represented by a black square. The N-terminal NPL and C-terminal FKBD domains are represented by the blue and green triangles, respectively, and the central highly charged region is represented by an orange triangle which is clearly shifted towards the region occupied by IDPs. (C) Structure of NPL domain of *Drosophila* Fkbp39 (at the N-terminal) generated using PyMOL (PDB: 4CA9; The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). (A) and (B) adapted from ref. 17.

Being comprised of 361 amino acid residues and weighing 39.3 kDa, *S. pombe* Fkbp39 is thought to be involved in nucleosome assembly, as inferred from sequence orthology (ISO) with *S. cerevisiae* Fpr3[19]. This proline isomerization influences H3 methylation, thus indicating a role in gene expression regulation. In *S. cerevisiae*, orthologs of the aforementioned PPIase are found, namely, Fpr3 and Fpr4[20,21]. Indeed, Fpr3 and Fpr4, which contain an acidic domain in addition to the conserved PPIase domain, have been shown to alter chromatin structure in a similar fashion to histone chaperones. Park, Xiao and Lei suggested that the acidic domain eases histone deposition and may be involved in rDNA silencing[22]. Fkbp39, like Fpr4, has a NPL domain, and studies in *D. melanogaster* have shown that this NPL domain located on the N-terminal side of the protein most possibly resembles a pentameric structure highly related to the nucleosplasmin domain, which is settled on a β-propeller architecture

consisting of a set of five monomers forming the stable pentameric structure[12]. Although this pentameric structure was suggested to occur on the N-terminal domain, a recent study from Koztowska et al., in opposition to previous declarations mostly based on studies of isolated NPL domains (via SANS, X-ray diffraction, and NMR), states that the NPL domain of the full-length Fkbp39 is possibly incompatible with the formation of pentameric complexes. Instead, their work reported that the NPL domain, in the full-length Fkbp39 in *D. melanogaster*, may exhibit a quaternary structure instead of the conserved pentameric structure of nucleoplasmin core domain and other proteins containing the NPL domain in the N-terminal domain, basing their statement on MALS, SE AUC, and SAXS (which revealed that oligomerization occurs via the NPL domain) techniques. They also pointed out the dynamic and flexible nature of Fkbp39, which is thought to be due to the existence of highly flexible disordered regions. Koztowska et al. Suggested that the tetrameric structure may play an important role in binding to evenly numbered histones during nucleosome assembly and chromatin structure modification, highlighting that, despite the high sequence identity between nucleoplasmin core domains, structural architecture can vary significantly. Nevertheless, they recommend interpreting their results with caution as NPL and nucleoplasmin core domains might behave differently when explicitly isolated, or in the full-length protein context. Moreover, they also propose that the differential observation of structural organization in each of the aforementioned contexts might be associated with the presence of alternative oligomerization interfaces (e.g. a long disordered linker)[17]. Previous work by Ramos et al. showed that the disordered distal region of nucleoplasmin from *X. laevis* promotes the recognition of distinct oligomerization states of histones[23].Thus, Koztowska et al. advanced that the flexible linker region found in Fkbp39 might fine-tune the tetramer structure and the oligomerization potential of the NPL domain[17]. Furthermore, this NPL domain is also capable of associating with histones, both in vivo and in vitro, albeit less intensely than in vivo. Edlich-Muth et al. suggested that Fkbp39 would hence interact with nucleosomes instead of histones. Noteworthily, Fkbp39 has been shown to co-purify with kinetochore proteins, as well as with microtubules. Moreover, its NPL domain is capable of binding divalent metal cations, suggesting that Fkbp39 harbours catalytic capabilities. In yeast Fpr4, binding to H2B nuclear localization sequence and regulation of lysine methylation (and consequently, gene expression), inter alia, has also been reported. A significant number of Fpr4´s characteristics are effectively shared with Fkbp39, emphasizing the idea that it may be a functional homologue of Frp4. Edlich-Muth et al. also showed that Fpr4 could form a pentamer in vivo. This is a valuable discovery, since it would then imply that binding of Fpr4 (and hypothetically, by analogy, of Fkbp39) to histones would position the five PPI domains effectively close to histones and neighbouring proteins, which would be reflected in a possibly bolstered local PPIase activity[12]. Further research is needed to elucidate Fkbp39's role and potential protein-protein interactions. To that intent, deeper insight into each of the composing domains of the protein is required. A particular focus on the understanding of the pentamer-to-tetramer transition (from isolated NPL core domains to NPL domains in the full-length Fkbp39) is of utmost importance to gather additional information in regard to a potential mechanism of action of nuclear Fkbps[17]. Solving the structure of the full-length Fkbp39 would be a very good kickstart to attempt to understand how the protein really works.

1.2.5.2. Line-of-thought and trends

Currently, in silico identification of histone chaperones remains challenging due to their diverse structural nature. It is thought that acidic patches present in many histone chaperones have a preponderant role in stabilizing the electrostatic interactions with the positively charged histones[4]. Current evidence points out to the "shielding" ability of histone chaperones which can protect histones from unspecific interactions through their binding to charged histone-DNA and histone-histone contact sites. Moreover, they can also facilitate interactions between histones and DNA by forming less stable histone-chaperone intermediates, thus playing an important role in chromatin remodelling[24].

1.2.6. Outlook

As we are uncovering the existence of various enzymes able to catalyse the modification of amino acid residues located in the globular region of core histones, the predominant line-of-thought which was centred on the notion that PTMs occurred almost entirely on the histones tails has been shifting. Nowadays, the scientific community is looking at PTMs, histone mutations and DNA sequence variations to study their effect on nucleosome mobility and chromatin structure formation. Although this has been a tough task for multiple reasons: 1) the highly dynamic and flexible nucleosome can only be studied through "static snapshots"; 2) there is a lack of structural data for higher-order chromatin organization; and 3) in vitro models of higher-order chromatin structures are experimentally challenging to obtain. Taking the aforesaid in consideration, there is still a long way ahead to uncover the totality of the biological role played by the nucleosome[25].

1.3. Nucleosome-interacting proteins

As mentioned beforehand, genomic DNA can be accessed through the action of chromatin-remodelers. There are two major types of chromatin remodelling complexes: 1) covalent histone-modifying complexes; and 2) ATP-dependent chromatin remodelling complexes. These protein complexes play crucial roles in chromatin deposition and assembly, chromatin editing, and chromatin accessibility.

Covalent histone-modifying complexes can catalyse the addition or removal of chemical groups through various enzymatic modifications (e.g. acetylation, methylation, phosphorylation, and ubiquitination), these occur in the majority of cases at the N-terminal histone tails (thought to be unstructured in the context of a single nucleosome). Direct consequences of these modifications are reflected on the binding affinity between the histones and DNA.

Most importantly, ATP-dependent chromatin remodelling complexes, which utilize the energy of ATP hydrolysis to function, play a central role in the assembly of chromatin, the accessibility of the most diverse factors to chromatin, as well as the restructuring of nucleosomes. These protein complexes do so by packaging and unpackaging chromatin, exposing parts of the genome which must be properly and tightly regulated in order to have control over cellular processes such as gene transcription, DNA replication, repair and recombination. ATP-dependent chromatin remodelling complexes are classified according to their role. Thus, three main types can be described based on ATP-dependent chromatin remodelling complexes involved in: 1) chromatin assembly (the majority of ISWI- and CHD-family members); 2) chromatin editing (mainly the INO80/SWR1 family which is involved in histone exchange through removal or replacement of noncanonical histone variants); and 3) chromatin accessibility (mainly the SWI/SNF family responsible for the repositioning and ejecting of the histone octamer, eviction of dimers, and exposure of sites susceptible to be bound by DNA-binding proteins)[4].

1.4. The histone code hypothesis & the role of histone core distortion

Nowadays, the consensus points to the view that changes in gene expression can occur without changes in DNA sequence (i.e. epigenetics). Considering the influence of the vast array of PTMs on histones which had the most diverse and complex consequences over the accessibility of the underlying DNA, a new concept arose: "the histone code"[26–30], which supported the hypothesis that the presence of different PTMs on histones could lead to varying binding affinities for the various proteins with the ability to bind to chromatin, subsequently influencing transcriptional states. Notably, this hypothesis can be seen as incremental in the sense that it can further extend the information potential encoded in the genome, reminding us of the importance and relevance of the role played by epigenetics in mechanisms involving chromatin[31,32]. These modifications (e.g. acetylation and methylation) occur mainly on histone tails (see Figure 1.6) and are catalysed by enzymes which show a high specificity for distinctive amino acid positions. Additionally, some modifications have shown to be dependent on others, further reinforcing the complexity of the histone code. To date, several domains have been shown to selectively recognize covalent modification marks (e.g. the bromodomain recognizes acetylated lysine residues, and the chromodomain specifically binds to methylated targets). These effector domains are commonly found in remodelers and multiple transcription factors. Acetylation and methylation marks can set off the recruitment of chromatin-modifying proteins. Additionally, the phosphorylation of tyrosine, serine, and threonine residues, as well as the ubiquitylation of lysine residues, can also be recognized by specific domains[3].

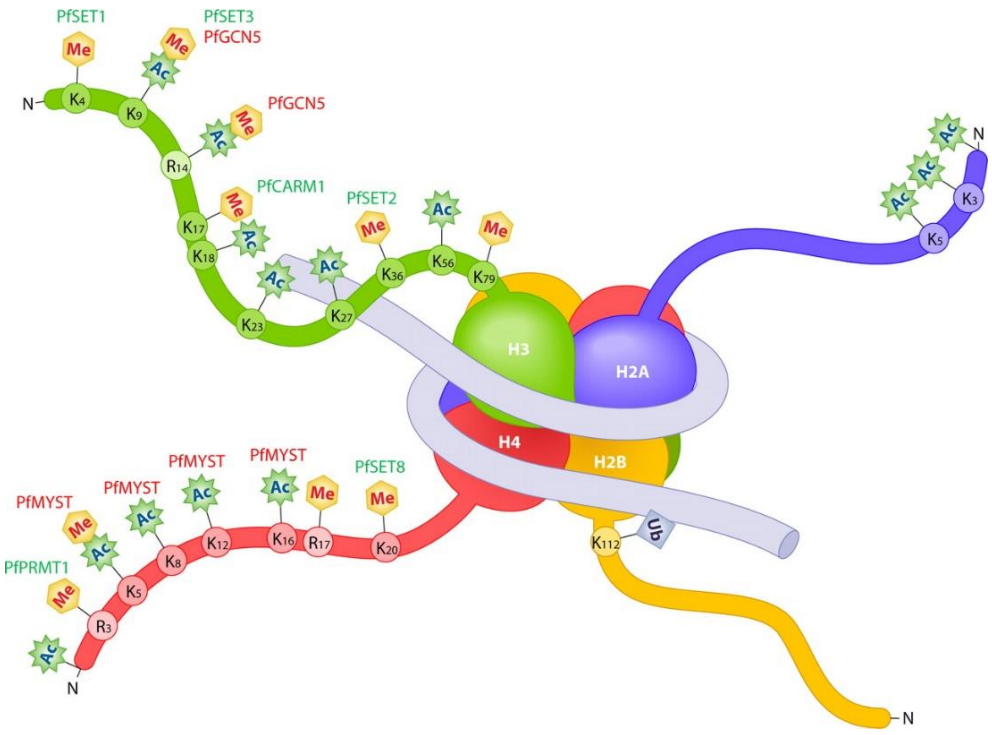


Figure 1.6 Graphical representation of a nucleosome exhibiting the four canonical core histones and multiple PTMs on the histones tails in *P. falciparum*. Each histone (H2A, H2B, H3, and H4) is coloured differentially. Amino acid residues where PTMs occur are represented by their respective single-letter code and linked to their known reported PTMs. Adapted from ref. 33.

More recently, following the discovery of counter examples for most of the histone crosstalks (e.g. H3K4 and H3K36 methylation which can recruit both histone acetyltransferases – HATs – as well as histone deacetylases – HDACs) which increases the complexity in attributing exact roles to each histone modification, chromatin research has shown an increasing focus on the role of the order and mechanism of addition and removal of these PTMs over the readout of a particular gene[34]. Furthermore, a study by Wang and colleagues has shown that recreating histone PTMs in vitro that influence gene expression, albeit possibly enabling the recruitment of RNA Pol II, was not sufficient for transcription to occur. Hence, focusing exclusively on mapping histone modification patterns is not enough, there needs to be an extensive understanding over the recruitment and regulation of the protein complexes involved in the addition and removal of these epigenetic marks. Only through the knowledge of the whole can we try to start unravelling the mechanisms that govern gene expression at this level. Also, with the rise of genome-wide experimental techniques in the 1990s, and the advances over its optimization in the last decade, it has been made possible to map histone modification marks, the enzymes which catalyse these PTMs, as well as the various factors involved in these reactions. Nowadays, it is clear that the regulation of gene expression is not solely governed by transcription factors, but also by the myriad of biological events that can alter histones´ composition. It appears the future of research focused on the histone code is shifted towards the uncovering of additional histone crosstalk networks and the role of order of appearance in each of these complex histone crosstalks[34].

Today, the notion that the histone octamer core is somewhat more "plastic" than previously thought is broadly accepted, providing histone core distortion with new-found roles in chromatin-templated processes[35].

1.4.1. Histone post-translational modifications

1.4.1.1. Acetylation

Histone acetylation was the first histone modification to be extensively described, being discovered in 1961, pioneering studies inferred its association to active transcription states. This was later in agreement with posterior research which indicated that acetylation neutralizes the positive charge of lysine residues, which directly conflicted with the electrostatic interactions localized at histone-DNA contacts through its weakening, leading to a subsequent increase in the accessibility of the underlying DNA, thus facilitating gene expression.

1.4.1.2. Methylation

Histone methylation, on the other hand, exhibits a lesser direct influence on nucleosome dynamics when compared to acetylation, which is inferred by the observation that mono-, di-, and trimethylation of lysine residues in histones have no effect on its positive charge. To date, multiple histone lysine methylations have been associated with either active (e.g. H3K4me), or repressed (e.g. H3K9me) transcription states, thus increasing the complexity of the histone code[31]. Histone methylation is thought to be the most diverse PTM occurring on histones in matters of the numbers of targeted residues, its potential for signalling, as well as its role in the regulation of multiple biological processes. Histone methylation is reversible and occurs on the N-terminal of lysine (K) and arginine (R) residues. Lysine residues can be mono-, di-, and trimethylated (denoted as "me1", "me2", and "me3", respectively), and arginine residues can be subjected to mono-, and dimethylation, although the effects of arginine methylation are less clearly understood. Following the typical nomenclature for histone methylation, methylated residues in histones are denoted with the respective histone, followed by the residue single-letter code, its position in the polypeptide chain, and the respective methylation state (e.g. "H3K79me2" indicates that the lysine residue at position 79 in core histone H3 is dimethylated). Methylation on the latter amino acid residues, albeit only faintly altering the primary structure, substantially increases the amount of encoded biological information within the molecule. In vivo, various canonical evolutionarily conserved lysine and arginine methylation sites are found (e.g. in humans, H3K4, H3K79, and H4R3). To note, some histidines (H) have been shown to be monomethylated, albeit being rarely identified. In the cellular context, the direct effect of methylation of proteins (performed by "writers", which catalyse the addition of a methyl group, and "erasers", which remove methyl groups) is reflected in the signalling responsible for regulating protein-protein interactions. To sum it up, methylation events that occur in histones can trigger a myriad of DNA-templated processes (e.g. activation and repression of transcription) through the action of the so-called "readers". Loss of control over histone methylation has a drastic impact on the health status of the affected, being linked to cancer, aging and disorders of varying magnitude[4].

Furthermore, it has been shown that histone methylation is present in both eu- and heterochromatin and can act as a long-term epigenetic mark. In the face of a non-removable histone lysine methylation imprint, an option for "camouflaging" this mark would be to "dilute" the histone lysine methylation levels through DNA replication and nucleosome distribution[31]. Until 2004, histone methylation was thought to be virtually irreversible. That same year, Shi et al. discovered the first demethylated histone ("LSD1", nowadays termed "KDM1A")[36], reviving interest among the scientific community. Today, two major classes of histone demethylases have been recorded and classified according to their mechanism of action.

1.4.1.3. Phosphorylation

It is common knowledge that phosphorylation provides a negative charge to the modified residue. This has direct implications over the strength of interactions between histones and DNA. In fact, phosphorylation of histones generates electrostatic repulsion towards the negatively charged DNA, weakening these interactions. In accordance with this, it has been observed that phosphorylation of H3T118 leads to a more accessible chromatin by interfering with nucleosome wrapping. Additionally, histone phosphorylation is known to happen in a myriad of biological processes, such as cellular response to DNA DSBs, role in development and alteration of the affinity of chromatin-binding proteins to their targets[30].

1.4.1.4. Ubiquitination

Although less clearly understood, studies have shown the role of ubiquitination in mediating gene silencing. In yeast, the ubiquitination of H2B regulates H3 methylation and impacts gene silencing[37]. This is a well-fitted example of "histone crosstalk" – also termed "trans-tail regulation of histone modifications", in the case of PTMs being implemented on the histone tails. Fundamentally, ubiquitination of histones consists in the attachment of a ubiquitin moiety. If more than one ubiquitin is added to its substrate, the process is then naturally termed "polyubiquitination", which is strongly associated with the proteasomal degradation of proteins. On the other hand, monoubiquitination exhibits different roles, such as the regulation of transcription for example. When comparing the occurrence of ubiquitination on histones and on the majority of other proteins susceptible to this PTM, histones exhibit predominantly monoubiquitination rather than polyubiquitination, core histones H2A and H2B being the most ubiquitinated among all histones[4].

1.5. Nucleosome mobility

Initially, the pseudo-symmetric nature of the nucleosome was thought to reflect tightly bound DNA at each histone-minor groove contacts leading to a robust molecular structure. However, ensuing single molecule experiments observing the energetics of histone-DNA contacts have shown that these vary extensively in the magnitude of strength between the central dyad, and entry and exit sites. Hence, nucleosomes are currently known to be dynamic and capable of sliding along DNA in vitro. In vivo, this sliding is made possible by the action of ATP-dependent nucleosome-remodelling complexes. In a broad view, these protein complexes are responsible for generating a remodelled nucleosome which is relatively stable and shows an increased mobility, facilitating the access of the underlying DNA to the myriad of nucleic acid-interacting proteins[3].

To date, several mechanistic models have been proposed in an attempt to expose how nucleosome mobility works. First and foremost, the "twist diffusion" (or "twisting") model, which states that single base pairs can be transferred between the linker DNA and the DNA that is coiled around the histone octamer core, allowing for the twisting or untwisting of this nucleosomal DNA which is then reflected in the loss or gain of base pairs. If propagation of the twist defect occurs along the DNA coiled around the nucleosome (i.e. "twist diffusion"), the histone core would then physically shift along the DNA. Although structural variability around the histone octamer core supports the latter model, the shifting of DNA outside the boundaries of the histone core would require an alteration in the rotational phasing of DNA linked to the histone octamer core. Hence, albeit a pertinent mechanism for explaining the shifting of nucleosomes along DNA in response to thermal fluctuations, it is more likely that chromatin-remodelers utilize a distinct mechanism for ATP-dependent sliding of nucleosomes[38].

The loop (or bulge) propagation model, similar to the twist diffusion model, is based on the premise that DNA from one linker is capable of transiently shifting onto the nucleosome, thus allowing for the formation of a loop of DNA which then diffuses around the histone core in a wave-like propagation. To note, this model differs from the latter in the proposal that DNA is shifted onto the nucleosome in larger segments (i.e. loop propagation) which, although more expensive in energetical terms, compensates by allowing for the maintaining of DNA rotational phasing on the histone core, unlike the twist diffusion model[38].

Alternative models (and variations of pre-existing models) for nucleosome mobility have been proposed. Notably the histone core swivelling model, which proposes that a remodeler-induced change in the structure of DNA results in the modification of the local landscape of histone-DNA contacts. These structural changes favour the shifting towards the chromatin-remodeler through the swivelling of the histone core. This swivelling allows for the shifting of the histone core by one minor groove, thus physically translocating the nucleosome along DNA while maintaining DNA rotational phasing[38].

To note, several studies suggest that histone modifications on the nucleosome surface play a crucial role in modulating histone-DNA interactions, thus influencing the stability of nucleosome positioning and the efficiency of nucleosome translocation. Based on the latter premise, an elegant model has been proposed by Cosgrove and colleagues, the "regulated nucleosome mobility" model[3]. This model suggests that histone side chains proximal to histone-DNA contacts can be modified by the concerted action of chromatin-remodelers and histone-modifying enzymes independently of the presence or absence of nucleosomal DNA. Through these modifications, histone-DNA interactions are altered and lead to a modulation in the accessibility of genomic DNA as well as in its higher-order packaging into chromatin[3]. To support this model, several unbiased genetic screens have been performed in yeast. To sum it up, there is a strong consensus that nucleosome mobility, with its roles in nucleosome

disruption and repositioning, is a major component of chromatin remodelling processes during transcription[11].

1.5.1.Nucleosome phasing

Naturally, histones are known to bind to a wide variety of DNA sequences. Despite that, past studies have shown a preferential biding of the histone octamer to a specific sequence, termed "Widom 601"[39]. This observation has been coined "nucleosome phasing", i.e. the directing of the histone octamer to bind preferentially to a defined sequence. Later, other strong-positioning DNA sequences have been identified, e.g. the 3′-LTR of the mouse mammary tumour virus (MMTV-A)[40]. Nucleosome phasing was a perplexing concept because histones are viewed as not sequence-specific proteins. But the answer lies in DNA itself: DNA sequence dictates its structure and bendability. That is, a defined DNA sequence that can easily bend and mould the nucleosome would lead to a substantial decrease in the free energy of binding. It is also known that the main parameter of nucleosome positioning is derived from its rotational setting which is intrinsically related to the ability of a particular DNA sequence to coil around the histone octamer. Later experiments with designed DNA sequences that would replicate the ideal sequence for nucleosome phasing have successfully shown that the rotational setting truly played a major role in nucleosome positioning[11]. These strong-positioning DNA sequences can be used to assemble nucleosomes in vitro.

1.5.2.Nucleosome stability

When addressing the concept of nucleosome stability, one must tackle with the following question: "what determines nucleosome stability?". There are multiple paths one can take to try to answer this question. Some studies have focused on the analysis of non-nucleosomal states, other projects discussed DNA accessibility as a determining factor of nucleosome stability. Andrews and Luger aimed at the thermodynamics of nucleosome (dis)assembly. A direct approach to measure nucleosome stability would clearly be the dilution of preformed nucleosomes to identify the fraction of free DNA against nucleosomal DNA. The only problem with this approach resides in the fact that the reaction is not at equilibrium due to nucleosome assembly being impaired because of high dilution conditions. And, to acquire any thermodynamic data, equilibrium is a requirement[41].

In *X. laevis*, reconstituted NCPs populations are subdivided into the asy- and symmetrically positioned histone octamers in regard to the 146 bp DNA. Incubation at higher temperatures yields a higher amount of the end-positioned NCP – nucleosomes which shift toward the DNA ends. The number of histone-DNA contacts and the DNA sequence itself influence the temperature and time required for the complete shifting to end-positioned species[25].

1.5.3.Effect of temperature on nucleosome mobility

Nucleosomes tend to be highly resilient to thermal and salt-induced dissociation. This is thought to be due to the high number of existing histone-DNA contacts which reinforce the overall structure integrity[42]. Nevertheless, previous research has shown that nucleosomes containing short nucleosomal DNA exhibit shifting when incubated at high temperatures for some time. Specifically, nucleosomes seem to shift towards the DNA ends when submitted to high temperatures[42] (see Figure 1.7 A), which could potentially alter chromatin accessibility and could be explained by the energetic penalty of DNA crossover at the entry and exit sites of nucleosomes. Flaus and Owen-Hughes showed that when the linker DNA length increases, the incubation temperature required for observing nucleosome shifting decreases[42] (see Figure 1.7 B). It is not yet fully understood why nucleosomes tend to shift to DNA ends when incubated at higher temperatures. Flaus and Owen-Hughes argued that it would most possibly not be so relevant in vivo since only a small number of end-positioned nucleosomes are observed under those conditions. Additionally, they showed that the nucleosome was more susceptible to shifting when

a long DNA arm was present and that a fraction of the nucleosomes could move up to 4 bp beyond the DNA end[42]. Furthermore, although no data was shown to back it up, they noticed that nucleosomes possibly reached equilibrium since they were unable to detect modifications in the nucleosome positioning when incubating for up to 8 hours.

Noteworthily, Flaus and Owen-Hughes found that nucleosomes were shifted in a stepwise manner with increments of 40 to 60 bp DNA. Hence, in theory, DNA accessibility to external factors is altered to a greater extent than it would be in the case of single-base increments in nucleosome shifting. They proposed that this would allow the amplification of the effects of signalling pathways at promoters, and that it might influence the dislocation of neighbouring nucleosomes in vivo, affecting the higher-order chromatin structure[42]. Also, they noticed that nucleosomes exhibit different preferred positions when comparing their shifting during chromatin assembly and during thermal-induced repositioning, which would then support the assumption that the sites preferred by nucleosomes in the chromatin assembly context are different from those preferred by intact nucleosomes[42]. It might be that the different nucleosome states (e.g. tetramer, hexamer, etc.) have different preferred positions.

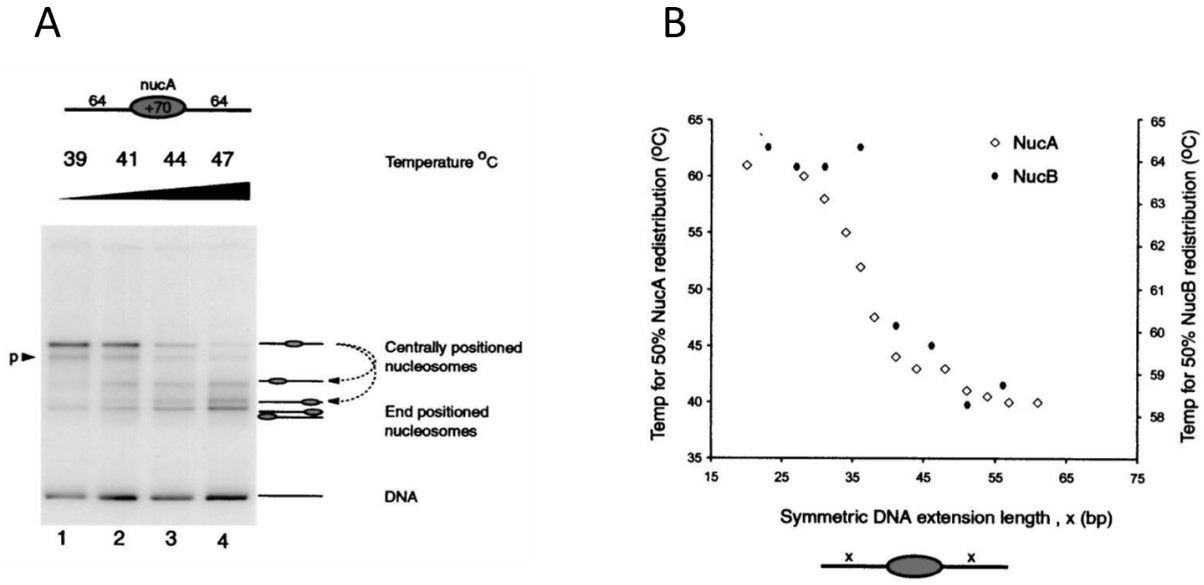


Figure 1.7 Shifting behaviour of MMTV NCP species. (A) Thermal-induced shifting of nucleosomes. (B) Influence of linker DNA length on the required incubation temperature for nucleosome shifting. Nucleosomes with asy- and symmetrically positioned histone octamers in regard to DNA are depicted accordingly. Incubation temperature set at 39, 41, 44, and 47 °C. "p": intermediate species. Adapted from ref. 42.

1.6. State-of-the-art

To reconstitute chromatin in vitro, as extensively addressed by Lusser and Kadonaga, two paths exist: 1) the ATP-independent random deposition of histones onto DNA, which requires core histones, DNA and a histone "chaperone" (e.g. NaCl); and 2) the ATP-dependent assembly of nucleosome arrays containing periodically positioned nucleosomes, requiring chromatin assembly factors (e.g. CAF-1) which are subdivided in ACFs and RSFs[43,44]. Different methods for chromatin reconstitution exist to this day, as depicted in Figure 1.8. Depending on the type of assays to be executed downstream, mononucleosomes – a nucleosome monomer – or nucleosome arrays might be used. With the latter, one can choose to reconstitute periodically or randomly distributed nucleosomal arrays. To obtain native core particles, simply digesting chromatin with micrococcal nuclease is enough. If one wants to generate mononucleosomes (usually harbouring more than 147 bp and less than 250 bp of DNA), purified core histones and DNA can be used (this is the approach used in the current project). DNA with varying lengths affect nucleosome reconstitution differently[45]. Usually, the reconstitution is performed via salt dialysis with NaCl. For studying the binding of proteins to nucleosomes, mononucleosomes and core particles can be used. Additionally, mononucleosomes can be used for screening the activities of ACFs (e.g. role in nucleosome sliding) to check for potential histone-modifying enzymes and for structural biology assays (e.g. X-ray crystallography). Resorting to mononucleosomes for chromatin studies has multiple advantages: 1) reconstitution is well-characterized and uses purified core histones and defined DNA sequences; 2) gel-shift analyses and DNase I footprinting allows for the detection of binding proteins; 3) nucleosomes can be specifically positioned by using well-defined DNA sequences with distinct structural characteristics; 4) native-PAGE allows the detection of differentially positioned nucleosomes (position of histone octamers in regard to DNA) and the monitoring of nucleosome sliding. On the other hand, mononucleosomes and native core particles cannot fully reflect chromatin in vivo[43]. Noteworthily, it is imperative to know the native chromatin state in the experimental organism, i.e. chromatin structural organization varies between species (e.g. length of linker DNA). Knowing the native state of chromatin in the experimental organism allows for better experimental design later on[43].

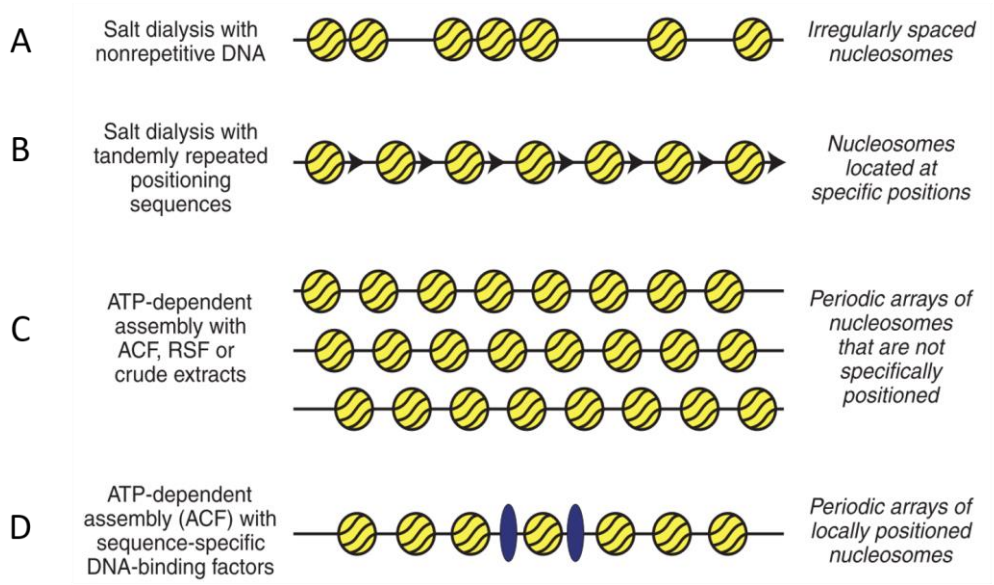


Figure 1.8 Methods for chromatin reconstitution in vitro using purified core histones and DNA. (A) Chromatin with nonrepetitive DNA reconstitution using salt dialysis; (B) Chromatin with tandemly repeated nucleosomes positioning sequences; (C) Chromatin with periodic nucleosome arrays reconstitution using ACFs, RSFs, or crude extracts; (D) Chromatin with specifically positioned nucleosomes using ACFs with sequence-specific DNA-binding factors (in blue). Adapted from ref. 43.

As mentioned beforehand, histones exhibit attractive hydrophobic and repulsive ionic interactions. When histones are kept in a high salt (e.g. 2 M NaCl) concentration environment, histone octamer core formation occurs due to the hindering of the repulsive ionic interactions. Conversely, in low salt concentration conditions, histones mostly assemble in (H2A-H2B) dimers and $(H3\text{-}H4)_2$ tetramers[43,46,47]. For core histones reconstitution, both with WT and variant species, various experimental approaches have been previously described[43,48,57,49–56]. One great advantage of resorting to recombinant histones for chromatin studies is the absence of eukaryotic factors which usually copurify with histones during the purification process, the apparent absence of PTMs and the ability to generate diverse mutant variants[43]. Furthermore, Lusser and Kadonaga also stated that histones typically stick to plastic and glass surfaces, thus recommending the use of siliconized materials. When running an SDS-PAGE, from the slowest to the fastest, core histones migrate in the following order: H3, H2B, H2A, and H4. A good indicator of the quality of the purification is denoted by the presence of full-length H3. H3 is the most proteolytically sensitive core histone, hence being more susceptible to degradation. If H3 degrades along the purification method, a band appears on SDS-PAGE gels between H2A and H4[43]. Core histones can be purified both from inclusion bodies – intracellular aggregates usually made of proteins – or directly using whole cell extracts under strong denaturing conditions that allow the solubilization of inclusion bodies[57]. An SDS-PAGE gel allowing for the comparison of these two methods is shown below in Figure 1.9.

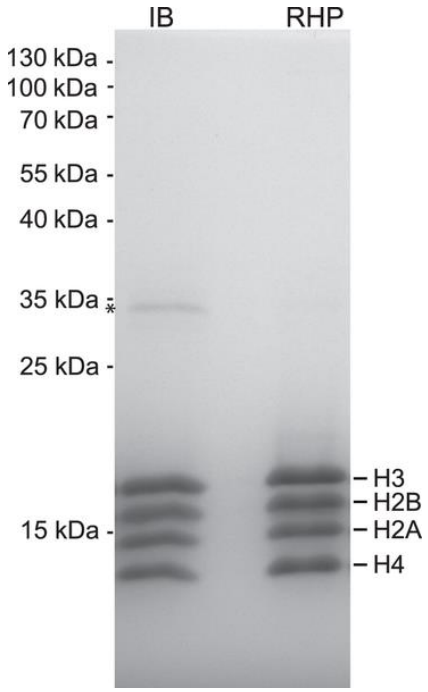


Figure 1.9 SDS-PAGE gel exemplifying two methods for histone purification. IB: Inclusion bodies purification; RHP: Rapid histone purification (as proposed in ref. 57); *: Contamination. Each core histone (from *Drosophila*) is marked accordingly. Adapted from ref. 57.

For this project, we followed an experimental procedure adapted from the nucleosome reconstitution methods described in refs. 54 and 57. A step by step diagram of the nucleosome reconstitution workflow used in this work is shown in Figure 1.10.

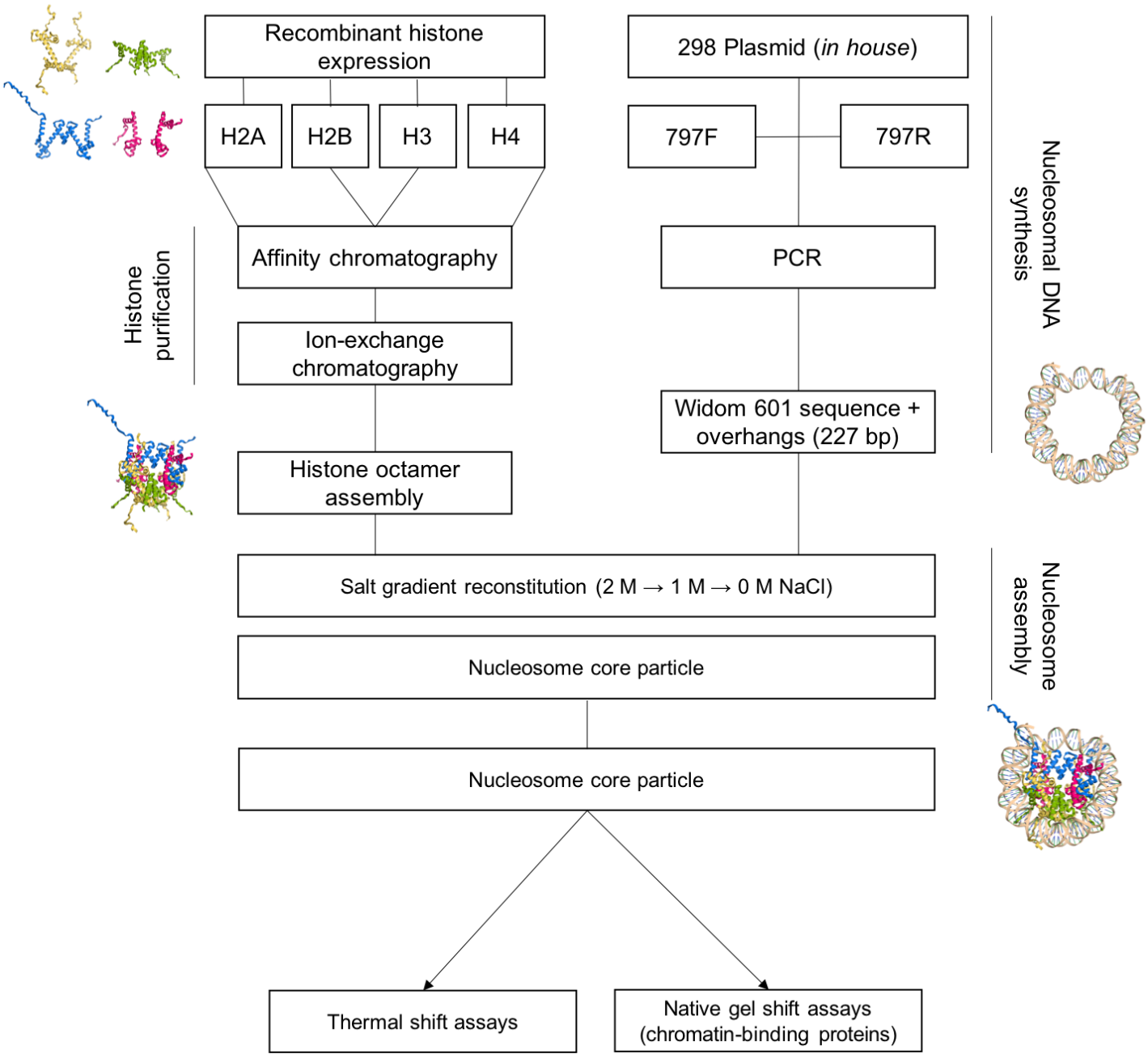


Figure 1.10 Nucleosome reconstitution experimental workflow. Core histones H2A in yellow, H2B in green, H3 in blue, and H4 in magenta. Nucleosomal DNA coloured in beige (helices only). Exemplifying for the long version of the Widom 601 sequence (227bp). Representations generated using PyMOL (PDB: 1AOI; The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

2. Aims and objectives

The study of the nucleosome, the fundamental unit of chromatin whose crystal structure at atomic resolution was solved more than two decades ago, is of vital importance for a better understanding of the regulation and storage of the extensive genetic information encoded in the genome. The discovery of its dynamic nature led to the assumption that it could act as an additional layer of gene expression regulation mechanism. In this work, we aim to observe and corroborate the previously described thermally-driven mobility of the nucleosome[42] using WT nucleosomes assembled from *X. laevis* core histones heterologously-expressed in *E. coli* and an extended version of the strong positioning Widom 601 sequence (227 bp)[39]. Furthermore, we also aim to study the mobility of modified nucleosomes harbouring different histone mutants using native gel shift assays. For each mutant, specific residues were replaced with cysteine residues. Moreover, for each modification in one specific histone, a complementary mutation in the other histone from the histone dimer pair will be introduced. These mutations will be designed using the 2.8 Å solved structure of the nucleosome core particle taking into account that the designed mutations are structurally close to each other to allow for the formation of disulphide bonds. The main aim of this project is to describe the influence of histone octamer distortion on the thermally-driven nucleosome mobility using a designed mutant nucleosome assembled from scratch.

Additionally, we also aim to observe the ability of Fkbp39, a previously described putative histone chaperone[58], to bind to *X. laevis* core histone dimer (H2A-H2B), and *X. laevis* core histone tetramer $(H3\text{-}H4)_2$ and observe the putative complexes via electron microscopy.

Lastly, we aim to screen the binding affinity of several chromatin-modifying enzymes to WT *X. laevis* nucleosomes via native gel shift assays.

3. Materials

3.1. Buffers

Table 3.1 List of buffers.

Buffer	Composition	Application
4x SDS-PAGE loading buffer	200 mM Tris-HCl pH 6.8 8% SDS 40% Glycerol 400 mM DTT	SDS-PAGE electrophoresis
4x separating gel buffer	1.5M Tris-HCl pH 8.8, 0.4% SDS	SDS-PAGE electrophoresis
4x stacking gel buffer	0.5M Tris-HCl pH 6.8, 0.4% SDS	SDS-PAGE electrophoresis
His buffer A	25 mM Tris-HCl pH 7.5, 1 mM DTT	Protein purification
His buffer B	25 mM Tris-HCl pH 7.5, 1 mM DTT, 2 M NaCl	Protein purification
His elution buffer	50 mM NaP pH 8.0, 300 mM imidazole pH 8.0, 2 M NaCl	Protein purification
His lysis buffer	50 mM NaP pH 8.0, 20 mM imidazole pH 8.0, 2 M NaCl, 3 μM β-mercaptoethanol	Protein purification
His washing buffer	50 mM NaP pH 8.0, 40 mM imidazole pH 8.0, 2 M NaCl	Protein purification
Histone dialysis buffer	25 mM Tris-HCl pH 7.5-8.0, 2 M NaCl, 1 mM DTT, 0.1 mM EDTA	Protein purification
Inclusion bodies washing buffer (IBWB)	25 mM Tris-HCl pH 7.5, 200 mM NaCl, 1mM DTT	Protein purification
Inclusion bodies washing buffer with triton (IBWB+T)	25 mM Tris-HCl pH 7.5, 200 mM NaCl, 1mM DTT, 1% Triton	Protein purification
Nucleosome dialysis buffer 0 M salt	15 mM Tris-HCl pH 7.5, 1 mM DTT	Nucleosome assembly
Nucleosome dialysis buffer 1 M salt	1 M NaCl, 15 mM Tris-HCl pH 7.5, 1 mM DTT	Nucleosome assembly
Protein dialysis buffer 150	30 mM Tris-HCl pH 7.5, 1 mM DTT, 150 mM NaCl	Protein purification
Protein dialysis buffer 200	30 mM Tris-HCl pH 7.5, 1 mM DTT, 200 mM NaCl	Protein purification
Protein dialysis buffer 200+1G	30 mM Tris-HCl pH 7.5, 1 mM DTT, 200 mM NaCl, 1% glycerol	Protein purification
Protein dialysis buffer 200+5G	30 mM Tris-HCl pH 7.5, 1 mM DTT, 200 mM NaCl, 5% glycerol	Protein purification
Protein dialysis buffer 200+10G	30 mM Tris-HCl pH 7.5, 1 mM DTT, 200 mM NaCl, 10% glycerol	Protein purification
Protein dialysis buffer 300	30 mM Tris-HCl pH 7.5, 1 mM DTT, 300 mM NaCl	Protein purification
Protein dialysis buffer 500	30 mM Tris-HCl pH 7.5, 1 mM DTT, 500 mM NaCl	Protein purification
S_{1000} buffer	25 mM Tris-HCl pH 7.5, 1 mM DTT, 2 M NaCl	Protein purification
S_{2000} buffer	25 mM Tris-HCl pH 7.5, 1 mM DTT, 2 M NaCl	Protein purification
S_{500} buffer	25 mM Tris-HCl pH 7.5, 1 mM DTT, 500 mM NaCl	Protein purification
S_{750} buffer	25 mM Tris-HCl pH 7.5, 1 mM DTT, 750 mM NaCl	Protein purification
TAE	9 mM Tris pH 7.6, 89 mM boric acid, 2 mM EDTA	Agarose gel electrophoresis
TBE	40 mM Tris pH 7.6, 20 mM acetic acid 20 mM acetic acid	SDS-PAGE electrophoresis

3.2. Antibiotics

Table 3.2 List of antibiotics.

Antibiotic	Application
Ampicillin	Bacterial transformation
Chloramphenicol	Bacterial transformation
Kanamycin	Bacterial transformation

3.3. Media

Table 3.3 List of media.

Medium	Composition	Application
LB	10 g/L Bacto tryptone, 10 g/L NaCl, 5 g/L Bacto yeast extract	Cell culture

3.4. Plasmids

Table 3.4 List of plasmids.

Plasmid	Insert
PET3a (#298)	Widom 601 sequence
pBlueScript	N/A
pPSU1	Penn State University DNA Ladder 1
pPSU2	Penn State University DNA Ladder 1

3.5. Primers

Table 3.5 List of primers.

Primer	Orientation	Sequence	Length (bp)	GC Content (%)	T_m (ºC)
1128R	Reverse	5'- CCC TGA TTT CTT CAG CTG -3'	18	50.0	48.0
1175F	Forward	5'- AGA TCT GTT TAG CTT GCC TCG TCC -3'	24	50.0	57.4
1175R	Reverse	5'- CAC CGT GTA TGA AAT CTA ACA ATG CG -3'	26	42.3	56.4
1182F	Forward	5'- GTA AAC GTC ACC GTA AAG TT -3'	20	40.0	47.7
1182R	Reverse	5'- AAC CAC CTT TAC CCA GA -3'	17	47.1	44.6
1183F	Forward	5'- GTC TGC GTG ACA ACA TC -3'	17	52.9	47.1
1183R	Reverse	5'- ATT TAC GGT GAC GTT TAG C -3'	19	42.1	46.8
1186F	Forward	5'- ACA GGA TGT ATA TAT CTG ACA G -3'	22	36.4	49.2
1229F	Forward	5'- TCG GTA CCC GGG GAT -3'	15	66.7	47.4
1248F	Forward	5'- TTG ACA CCC CGT GAC AGG ATG TAT ATA TCT GAC -3'	33	45.5	63.2
1249R	Reverse	5'- GTA TCA CGG GGT GAG ATC GCT CTG GAG AAT CCC GGT -3'	36	58.3	70.1
1250R	Reverse	5'- GTA TCA CGG GGT GAG ATC TCT GGA GAA TCC CGG T -3'	34	55.9	68
318F	Forward	5'- GCT AGT TAT TGC TCA GCG G -3'	19	52.6	51.1
423F	Forward	5'- GCA CAG GAT GTA TAT ATC TG -3'	20	40.0	47.7
423R	Reverse	5'- CTG GAG AAT CCC GGT -3'	15	60.0	47.7
438F	Forward	5'- CAG GGG CCC GGA TCC ATG TCA GGA AGA GGC -3'	30	66.7	69.8
439R	Reverse	5'- GAA CAG AAC TTC CAG CTC GAG GCT GCC GCG C -3'	31	64.5	69.7
596F	Forward	5'- CAG GAA ACA GCT ATG ACC ATG -3'	21	47.6	52.4
596R	Reverse	5'- TGT AAA ACG ACG GCC AGT -3'	18	50.0	48.0
797F	Forward	5'- TAT CCG ACT GGC ACC G -3'	16	62.5	48.5

797R	Reverse	5'- GAG TTC ATC CCT TAT GTG AT -3'	20	40.0	47.7
823R	Reverse	5'- GGA TTT GTA GCC TCC GTG -3'	18	55.6	50.3
872F	Forward	5'- TAA TAC GAC TCA CTA TAG -3'	18	33.3	41.2

3.6. Gels

Table 3.6 Composition of Native-PAGE and SDS-PAGE gels.

Type	Composition	Final volume (mL)
Native-PAGE	4 mL TBE 10x, 8 mL Acrylamide/BisAA 30% (37.5:1), 400 µL APS 10%, 40 µL TEMED, 27.5 mL H_2O	40 mL
SDS-PAGE	**Separating gel**: 7.5 mL 4x separating gel buffer, 17 mL Acrylamide/BisAA 30% (37.5:1), 300 µL APS 10%, 30 µL TEMED, 2 mL glycerol 25%, 5.5 mL H_2O **Stacking gel**: 3.75 4x stacking gel buffer, 2 mL Acrylamide/BisAA 30% (37.5:1), 150 µL APS 10%, 15 µL TEMED, 9.25 mL H_2O	30 mL

[1]

[1] β-mercaptoethanol and DTT have to be added to the respective buffers every three days since they tend to degrade with time.

4. Methods

4.1. Bacterial transformation of *E. coli* cells

Plasmid DNA was transformed into *E. coli* competent cells resorting to two different strains, namely, XL-1 Blue and BL21 (DE3) Rosetta. Firstly, 5 µL of plasmid DNA were added to 50 µL of competent cells in precooled 1.5 mL microcentrifuge tubes and carefully mixed, followed by an incubation on ice for 15 min. Next, the cells were heat shocked at 42 °C for 45s, and then put on ice for 2 min. 500 µL of LB medium were then added, and cells were put at 37 °C, 160 rpm for 1h in an orbital shaker to induce cell recovery, followed by plating on LB agar plates with the required antibiotics to allow for the selection of successfully transformed cells. In the case of XL-1 Blue, Kan was used, and for BL21 (DE3) Rosetta cells, both Kan and Cam were used. To allow cell growth, plated agar plates were incubated overnight at 37 °C. For histone expression, only the chemically competent BL21 (DE3) Rosetta cells were used.

4.2. Plasmid DNA purification

To obtain purified plasmid DNA, pelleted cells obtained from an overnight preculture grown at 37 °C in an orbital shaker (160 rpm) were treated following the NucleoSpin® Plasmid EasyPure protocol from Macherey-Nagel (http://www.mn-net.com). The washing step was performed twice to ensure yielding a "cleaner" product.

4.3. DNA sequencing

Firstly, bacterial precultures were prepared from single colonies in 5 mL of LB medium containing the appropriate antibiotic(s). Precultures were then put to grow overnight in an orbital shaker at 37 °C. The following day, precultures were centrifuged at 3900 rpm for 15 minutes to harvest cells, and the supernatant was discarded. Purified plasmid DNA was obtained resorting to the NucleoSpin® Plasmid EasyPure kit (as described in 4.2). Sequencing of DNA was performed via Sanger sequencing by GATC Biotech AG. For each sequencing job, a 20 µL aliquot of DNA or primers were sent. Primers were diluted to a final concentration of 0.1 µM. Sequencing results were analysed via FinchTV software (Geospiza, Inc.).

4.4. Polymerase chain reaction

The polymerase chain reaction was used for the amplification of the strong-positioning Widom 601 sequence (and variants) used downstream for nucleosome assembly (see subchapter 4.9). The PCR reaction is described below in Table 4.1.

Table 4.1 PCR reaction setup for the long Widom 601 sequence amplification (227 bp).

Reagent	Volume (µL)	Source
H_2O	33.5	In house
Phusion® HF Buffer (5x)	10	Phusion® HF Buffer Pack (New England Biolabs)
dNTPs (10 mM)	1	Jena Bioscience (Germany)
797F (Forward primer)	2	Metabion (Germany)
797R (Reverse primer)	2	Metabion (Germany)
298 (DNA Template)	1	In house
Phusion polymerase (in house)	0.5	In house

For the amplification of the long version of the Widom 601 sequence used in this work, initial denaturation of DNA was set at 95 °C for 30s, followed by 46 cycles of 10s at 95 °C (denaturation), 10s at 54 °C (annealing), and 10s at 72 °C (elongation). After the cycles ended, an additional step was set at 72 °C for 30s. Amplified DNA sequences were then analysed on 1% agarose gels (staining with DNA Gel Loading Dye 6x from Thermo Scientific).

4.5. DNA precipitation

Precipitation of DNA was performed by adding, first and foremost, 2.5 volumes of EtOH 100% and 10% of 3 M NaOAc pH 5.2. The sample was then let either 1h or overnight at -80 °C, followed by centrifugation at maximum rcf for 30 min. Supernatant was then removed and 500 µL of EtOH 75% was added to the DNA pellet. The ensuing solution was centrifuged at maximum rcf for 15 min and the supernatant was removed. The resulting sample was dried at RT for around 50 min, followed by the addition of 20 µL of H_2O and resuspended. DNA concentration was then assessed via the NanoPhotometer NP80 (from Implen), and its quality checked via 1% agarose gel (staining with DNA Gel Loading Dye 6x from Thermo Scientific).

4.6. Histone expression

Protein expression in the chemically competent BL21 (DE3) Rosetta cells was performed in LB medium. To that intent, when starting from glycerol stocks of the host laboratory stored at -80 °C, precultures (ranging from 5 mL to 30 mL) were firstly grown overnight in an orbital shaker at 160 rpm and 37 °C, and the respective antibiotic(s) were also added to the latter, as described in Table 4.2. Addition of cells from the glycerol stocks to their respective preculture recipient was done by picking some frozen cells using a 200 µL pipette tip. The following day, the cultured competent cells from the aforementioned precultures were harvested and used for growing larger cell cultures, with the addition of glucose 20%, as described in Table 4.2. These were grown at 37 °C, 150 to 200 rpm, until they reached an optical density at 600 nm (OD_{600}) of 0.6. Afterwards, temperature was lowered to 18 °C for 30 minutes, following the overnight induction of cells with 0.2 mM of IPTG. The following day, cells were centrifuged at 4000 rpm for 10 minutes and the pelleted cells were collected and either directly used for downstream applications, or stored at -20 °C.

Table 4.2 Components for cell induction.

Final volume of cell culture (LB, in L)	Added volume of glucose 20% (mL)	Added volume of antibiotic (mL)	Added volume of preculture (mL)
1	20	1	10
2	40	2	20

4.7. Recombinant protein purification from heterologous expression in *E. coli* BL21 (DE3) Rosetta strain

4.7.1. Affinity chromatography: 6xHis-tag purification (Ni-NTA Agarose)

Cells were resuspended in His lysis buffer using an Omnifix® Luer Solo syringe with Sterican® single-use hypodermic needles from Braun, this step is very important to avoid clogging the French press. Following the resuspension step, 200 µL of 0.1 M PMSF and 1 µL of DNAse (Pierce™ Universal Nuclease for Cell Lysis) were added. The next step consisted in the lysis of the treated cells resorting to a French press (25 ksi), followed by a centrifugation at 17000 rpm, 4 °C, for 15 minutes. The supernatant was then incubated with Ni Sepharose 6 Fast Flow histidine-tagged protein purification resin (from GE Healthcare) for exactly 30 minutes at 4 °C to allow for the binding of the histones to the resin. Next, the resin was washed 3 to 5 times with the His lysis buffer, being cautious about resuspending the resin each time, and the supernatant was discarded after each wash. The ensuing washed resin was then loaded onto a 10 mL disposable polypropylene column (from Thermo Scientific) at 4 °C. The resin was washed using 10 to 12 mL of His washing buffer, followed by the elution of bound histones using the His elution buffer. Elution efficiency of the proteins was analysed after running a 17 % SDS-PAGE gel and staining the gel with SimplyBlue™ Safe-stain. Lastly, the eluted histones were dialyzed overnight at 4 °C in the histone dialysis buffer.

4.7.2. Ion exchange chromatography: protein purification (Sepharose Fast Flow)

Following the overnight dialysis after the affinity chromatography, ion-exchange chromatography was performed using 10 mL disposable polypropylene columns from Thermo Scientific. If reusing columns that were properly washed and stored in EtOH 20%, firstly, the respective columns would be washed three consecutive times with ddH_2O, centrifuging at 3000 rpm, 4 °C, for 3-5 min. The washed resin would then be incubated and mixed with 5x His buffer A and the dialyzed proteins (the ones previously purified by affinity chromatography) at 4 °C for 30 minutes. After the binding of proteins to the resin, the sample was centrifuged at 2000-2500 rpm, 4 °C, for 4 minutes (supernatant, i.e. unbound protein, was collected for subsequent assessment of binding efficiency of the proteins to the sepharose resin). Remaining beads in the glass bottle were washed with S_{500} buffer. Next, at 4 °C, the resin was transferred to a S-column (using the S_{500} buffer to collect all the remaining beads). 2.5 volumes in column volume were then added for elution using S_{750}, S_{1000}, and S_{2000} buffers. Each fraction was collected individually. A 17% SDS-PAGE gel was run to analyse the efficiency of the purification and check for the enriched fraction(s). Aliquots of 0.5 mL to 1 mL were prepared from the enriched fractions, frozen in liquid nitrogen, and stored at –80 °C. To note, the used resin was thoroughly washed with 5 M NaCl, and 2 to 3 rounds of ddH_2O, and then stored at 4 °C in EtOH 20%.

4.8. Histone octamer assembly

To assemble the histone octamer, stoichiometric amounts of core histones H2A, H2B, H3 and H4 were mixed and incubated at 4 ºC overnight. Histone ratios were checked on 17% SimplyBlue™ Safe-stained SDS-PAGE gels prior to incubation.

4.9. Nucleosome assembly

As a means to assemble the nucleosome, equimolar mixing of the histone octamer and nucleosomal DNA must be observed. To that intent, in the current work, the long version of the strong-positioning Widom 601 sequence (227 bp) was used, as it has been previously shown to be a very strong and effective histone octamer binding sequence[39] (see subchapter 1.5.1). The salt concentration of the sample containing the nucleosomal DNA was adjusted to 2 M NaCl before mixing it with the histone octamer. Following the procedure for nucleosome reconstitution as described in ref. 54 (all steps performed at 4 ºC), each assembly sample was transferred in an Eppendorf lid covered with a pre-wetted 3.5 kDa cut-off standard cellulose regenerated membrane (Spectra/Por® 3 Dialysis Tubing). First and foremost, the sample was dialysed overnight in nucleosome dialysis buffer 1 M salt, followed by two rounds of dialyses in 0 M NaCl buffers. The first round of the dialysis in nucleosome dialysis buffer 0 M salt (final volume of 1 L) would take 5 to 6 hours (ideally, 5 hours), followed by the second round of dialysis in nucleosome dialysis buffer 0 M salt (final volume of 0.5 L), which would take 1 to 1h30 (ideally, 1h30). All three of these dialysis buffers were discarded after three uses. After the sample was thoroughly dialysed, nucleosome yielding was assessed by direct visualization on a native-PAGE gel right after staining with Invitrogen™ SYBR™ Gold Nucleic Acid Gel Stain (10000x concentrated in DMSO, from Fisher Scientific). This step allows to check for the presence of the nucleosome core particle, as well as the position of the Widom 601 sequence on the latter, and if present, free nucleosomal DNA. Successfully assembled nucleosomes were used later on for binding assays with multiple chromatin-binding proteins, as well for thermal shifting assays. Remaining nucleosomes were frozen in liquid nitrogen and stored at -80 °C.

4.10. Protein concentration

When required, proteins were concentrated to smaller volumes resorting to Amicon® Centrifugal Filter Units (from Merck), as described in Table 4.3, with varying MW cut-offs and volumes. Samples were added to the concentrating devices and centrifuged at 3900 rpm for 4 to 6 min. After each centrifugation, the input volume was pipetted up and down, being cautious not to touch the cellulose membrane with the pipette tip. Centrifugations were repeated until reaching the desired final volume, or if precipitation occurred (to avoid any further losses).

Table 4.3 Various devices Amicon® used throughout the project.

Device	Volume (mL)	Commercial name	Membrane cut-off (kDa)
Amicon® Ultra-4	4	Ultracel-30 regenerated cellulose membrane	30
Amicon® Ultra-15	15	Ultracel-30 regenerated cellulose membrane	30
Amicon® Ultra-15	15	Ultracel-50 regenerated cellulose membrane	50

4.11. *S. pombe* Fkbp39 purification and Fkbp39:$(H3\text{-}H4)_2$, Fkbp39:(H2A-H2B) complex formation assessment

As mentioned beforehand in subchapter 1.2.5.1, *S. pombe* Fkbp39's ability to form a complex with full-length (H2A-H2B) and $(H3\text{-}H4)_2$ from *X. laevis* was studied as a side project. To that intent, BL21 (DE3) Rosetta competent cells capable of expressing Fkbp39-SUMO-6xHis (*S. pombe*) were firstly grown in a total volume of 2 L (starting from a 30 mL preculture) as described in subchapter 4.6. Harvested cells were pelleted after centrifugation at 4000 rpm, 4 ºC for 10 min, and then resuspended. To resuspend pelleted cells, 90 mL of His lysis buffer was added and resuspended cells were transferred to a clean metallic cup using an Omnifix® Luer Solo syringe with Sterican® single-use hypodermic needles from Braun. Then, 100 µL of 0.1 M PMSF and 1 µL DNAse were added (for a final volume of 1L). Treated cells were then sonicated using a Branson Sonifier 250 (40% Duty cycle and 4 output control micro tip limit) for 10 min. Next, sonicated cells were purified by 6xHis-tag protein purification using Ni-NTA agarose, an affinity chromatography matrix used for purifying recombinant proteins carrying a 6xHis-tag, as described in subchapter 4.7.1. Elution fractions enriched in Fkbp39 were then collected, part of it was mixed with WT $(H3\text{-}H4)_2$ from *X. laevis* to check for the optimal ratio of complex formation, followed by the addition of PreScission enzyme (*in house*) in order to cut the SUMO-6xHis tag from Fkbp39. The resulting treated sample was dialysed using a 12-14 kDa cut-off standard cellulose regenerated membrane (Spectra/Por® 4 Dialysis Tubing) in the histone dialysis buffer overnight. The next day, the sample was transferred to a protein dialysis buffer 200 and dialysed for 5h. After the dialysis, the mixture was concentrated (using Amicon® Ultra-15, Ultracel-30 regenerated cellulose membrane, 30 kDa). The first attempt to concentrate the sample was unsuccessful due to the occurrence of precipitation and ensuing loss of sample. Hence, 10% glycerol was added to the mixture, and a new attempt to concentrate the sample was performed. This time, the sample was successfully concentrated to around 500 µL. The ensuing concentrated sample was run on a 17% SDS-PAGE gel to check if it yielded enough complex for downstream analyses, followed by size-exclusion chromatography and further concentration steps until obtaining pure enough complex.

The remaining volume of the elution fractions of Fkbp39 that was not mixed with histones was subdivided and dialysed overnight, after adding PreScission enzyme (*in house*), in three buffers with varying salt (NaCl) concentration (Protein dialysis buffers 150, 300 and 500) to check for optimal salt concentration in order to avoid Fkbp39 precipitation. 6-8 kDa cut-off standard cellulose regenerated membranes (Spectra/Por® 1 Dialysis Tubing) were used for these dialyses. The following day, the dialysed proteins were run on a 17% SDS-PAGE gel with the uncut protein to check for effective SUMO-tag removal. Fractions dialysed in 150 and 300 mM NaCl containing buffers were further purified through ion-exchange chromatography, whilst the 500 mM NaCl containing buffer fraction was stored at -80 °C. Enriched fractions from ion exchange chromatography of Fkbp39 were mixed in stoichiometric quantities with $(H3\text{-}H4)_2$ from the host laboratory after checking for the best ratio on a 17% SDS-PAGE gel. The ensuing mixture was then dialysed overnight in protein dialysis buffer 200+10G to avoid precipitation. Next, the dialysed complex was aliquoted (1 mL of sample in 2 mL polypropylene screw cap micro tubes, from Sarstedt) and centrifuged at 10k rpm, 4 °C, for 15 min. After centrifugation, soluble fractions were transferred to new screw cap tubes and the pellets were resuspended in 4x SDS-PAGE loading buffer. To check for any significant complex loss through precipitation, an additional 17% SDS-PAGE gel was run. Next, as the final aim of this side project was getting a stable and pure enough complex of Fkbp39 with $(H3\text{-}H4)_2$, and possibly (H2A-H2B), for downstream cryo-EM applications, a new dialysis was executed overnight in protein dialysis buffer 200, since it is not recommended to have glycerol in samples for EM analyses[59]. After mixing $(H3\text{-}H4)_2$ and Fkbp39 in a 1:1 ratio, 7 M guanidine·HCl was added to denature the proteins, and the mixture was

incubated at RT for 1h with slow mixing. The use of this strong chaotropic agent is justified by its ability to solubilize insoluble proteins (such as inclusion bodies), which is the first step to be taken for protein refolding experiments. After the incubation, the denaturated sample was centrifuged at 15k rpm, 4 °C, for 15 min. The ensuing soluble fraction was dialysed overnight in protein dialysis buffer 200+10G, allowing for the overnight refolding of the complex. Remaining purified Fkbp39 from before, as well as an aliquot of complex, were dialysed protein dialysis buffer 200 (without glycerol to check if Fkbp39 precipitates in buffer without glycerol when not in complex with histones, as well as in complex with histones). These three dialyses were executed using a 6-8 kDa cut-off standard cellulose regenerated membrane (Spectra/Por® 1 Dialysis Tubing). Then, ensuing dialysed samples were centrifuged at 15k rpm for 15 min at 4 °C, and aliquots of the soluble fractions were loaded on a 17% SDS-PAGE gel.

We also tried to check for complex formation of Fkbp39 with WT *X. laevis* (H2A-H2B) from the host laboratory. To that intent, (H2A-H2B) samples were loaded on a 17% SDS-PAGE gel with Fkbp39 to check for optimal ratio.

Due to the high incidence of precipitation either of the complex Fkbp39:$(H3\text{-}H4)_2$, either Fkbp39 alone in protein dialysis buffer 200+10G, new dialyses with varying salt concentration and glycerol percentage were prepared, as described in Table 4.4.

Table 4.4 Dialysis buffers' composition for Fkbp39:$(H3\text{-}H4)_2$ and FKbp39:(H2A-H2B) assembly.

Dialysed sample(s)	[NaCl] (mM)	% Glycerol
Fkbp39:(H3-H4) Fkbp39:(H2A-H2B)	200	1
Fkbp39:(H3-H4) Fkbp39:(H2A-H2B)	200	5
Fkbp39:(H2A-H2B)	200	10
Fkbp39:(H3-H4) Fkbp39:(H2A-H2B)	300	0
Fkbp39:(H3-H4) Fkbp39:(H2A-H2B)	500	0

Complex precipitation issue was ultimately resolved after using the aforementioned buffers. Hence, the complexes were aliquoted, part of these were frozen in liquid nitrogen and stored at -80 °C for later downstream applications (i.e. binding screening via gel mobility shift assays), and the remaining part was used for cryo-EM analyses at the Max Planck Institute of Biochemistry (Munich, Germany).

4.12. Chromatin-binding proteins native gel shift assays

Screening of chromatin-binding proteins binding to WT *X. laevis* nucleosomes was performed via native gel shift assays. Small volumes of concentrated WT nucleosomes (1 to 10 µL depending on nucleosome concentration) were mixed with small volumes of concentrated chromatin-binding protein (0.5 to 5 µL) and incubated at RT for 30 min or 1h at 4 ºC, followed by the addition of 1.5 µL of native-PAGE loading dye (40% glycerol and 2 µL of bromophenol blue in a final volume of 1 mL) and running the samples on 6% native-PAGE gels at 4 ºC. Staining of the gels was done with Invitrogen™ SYBR™ Gold Nucleic Acid Gel Stain (incubate twice for 10 min with gentle shaking, wash once with TAE 1x buffer).

4.13. Electrophoresis

4.13.1. Agarose

The lengths of plasmids, DNA, and the various nucleosomal DNA sequences prepared in this work were checked by visualization in agarose gels. All agarose gels were prepared the same way: 1% agarose gels run at 100 to 120 V for 20 to 30 min at RT. The ladder was prepared by mixing 4 µL of H_2O with 1 µL of DNA Gel Loading Dye 6x (Thermo Scientific) and 1 µL of GeneRuler 1kb ladder (Thermo Scientific). All samples were mixed using 5 µL of sample with 1 µL of DNA Gel Loading Dye 6x (Thermo Scientific). Visualization of the gels was done on a gel documentation system, adjusting the exposure appropriately. Each figure was printed using Mitsubishi K65HM-CE / KP65HM-CE HD thermal paper rolls.

4.13.2. Native-PAGE

Analysis of nucleosome assembly assays were done using Native-PAGE gels. For each sample, 10 µL of sample was mixed with 1.5 µL of native-PAGE loading dye, and the ladder was prepared by mixing 10 µL of H_2O, 1.5 µL of native-PAGE loading dye, and 1 µL of the GeneRuler 1kb ladder (Thermo Scientific). The gels were run at 200 V for 2h, at 4 °C. Afterwards, gels were carefully disassembled from the cassette and put in a thoroughly cleaned box shielded from light (the lid of the box was carefully wrapped in aluminium foil) containing 25 mL of 1x TAE buffer. Then, 2.5 µL of Invitrogen™ SYBR™ Gold Nucleic Acid Gel Stain (10000x concentrated in DMSO, from Fisher Scientific) was added to the solution contained the immersed gel, being careful not to pipette the stain directly over the gel but rather on the sides of the box in order to avoid any blob formation which would impair proper band(s) visualization. The box containing the gel was then put on a shaker at 50 rpm for 10 min at RT. Afterwards, the solution was discarded in the appropriate toxic waste disposal container, and 25 mL of 1x TAE buffer were added again to the box, putting the box to shake at 50 rpm for 10 min at RT, as before. The solution was discarded in the same toxic waste disposal container after the previous step, and the process was repeated once more. Lastly, the gels were analysed on a gel documentation system, adjusting the exposure appropriately. Each figure was printed using Mitsubishi K65HM-CE / KP65HM-CE HD thermal paper rolls.

4.13.3. SDS-PAGE

Protein purification was assessed through direct visualization of 17% SDS-PAGE gels after proper staining using SimplyBlue™ Safe Stain. For each sample, 10 µL of sample were mixed with 5 to 10 µL of 4x SDS-PAGE loading buffer, the ensuing solution was incubated at 90 °C for 3 min, and then loaded on the gel which was run at 300 V for 35 min. Excess SDS was removed after letting the gel in boiling deionised water thrice for around 30s, and staining was done by letting the gel stain in boiling SimplyBlue™ Safe Stain for a short period of about 10 to 20s. If required, destaining of the gel was done by letting the gel in boiling deionised water for around 30s, and repeated if needed.

5. Results

5.1. WT and mutant *X. laevis* histone purification

To avoid redundancy, this subsection will only cover two examples of histone purification, namely, *X. laevis* H4_V22C and WT H3, and H2Aub (ubiquitinated) and WT H2B expressed in BL21 (DE3) Rosetta cells Kan and Cam resistant (sample name: "1101"; see Table 5.1). The remaining histone purifications executed in the aim of this project are listed in Table 5.1. First and foremost, mutant and WT histones are purified from the pelleted cells as described in subchapter 4.7.1. All the histones purified in this project harbour a 6xHis-tag which allows for its purification via affinity chromatography. A 17% SimplyBlue™ Safe-stained SDS-PAGE gel was run to assess the purified histones yield. Usually, 10 µL of each fraction were loaded on the gel, using 4x SDS-PAGE loading buffer. 17% SimplyBlue™ Safe-stained SDS-PAGE gels for *X. laevis* H4_V22C and WT H3, and H2Aub and WT H2B purifications via affinity (A) and ion-exchange (B) chromatographies are shown in Figure 5.1 and Figure 5.2, respectively.

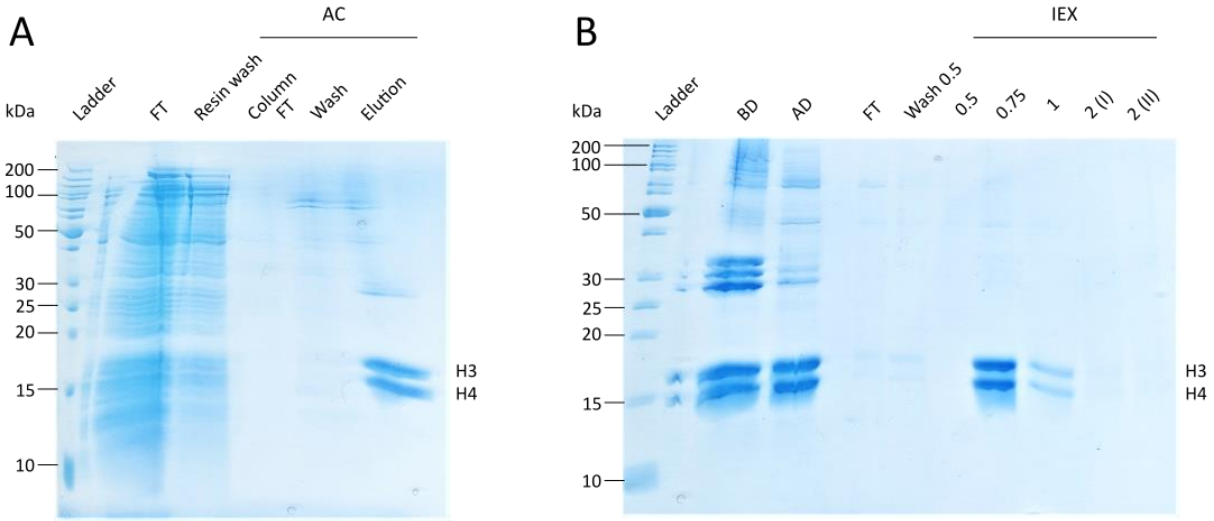


Figure 5.1 SimplyBlue™ Safe-stained 17% SDS-PAGE gels of 1101 (see Table 5.1) purification via affinity and ion-exchange chromatographies. (A) Affinity chromatography. (B) Ion-exchange chromatography. Ladder: PageRuler Unstained Protein Ladder (Thermo Scientific), FT: Flow-through; Resin wash: His lysis buffer wash; Column FT: Unbound protein after loading into the column; Wash: Washing step of the affinity chromatography; Elution: Elution step of the affinity chromatography; AC: Affinity chromatography; IEX: Ion-exchange chromatography; BD: Before overnight dialysis; AF: After overnight dialysis; Wash 0.5: Resin wash with 500 mM NaCl buffer; 0.5: 500 mM NaCl buffer; 0.75: 750 mM NaCl buffer; 1: 1 M NaCl buffer; 2 (I) and 2 (II): 2 M NaCl buffer (elution with 2 M NaCl buffer is executed twice). Core histones H3 and H4 are marked accordingly.

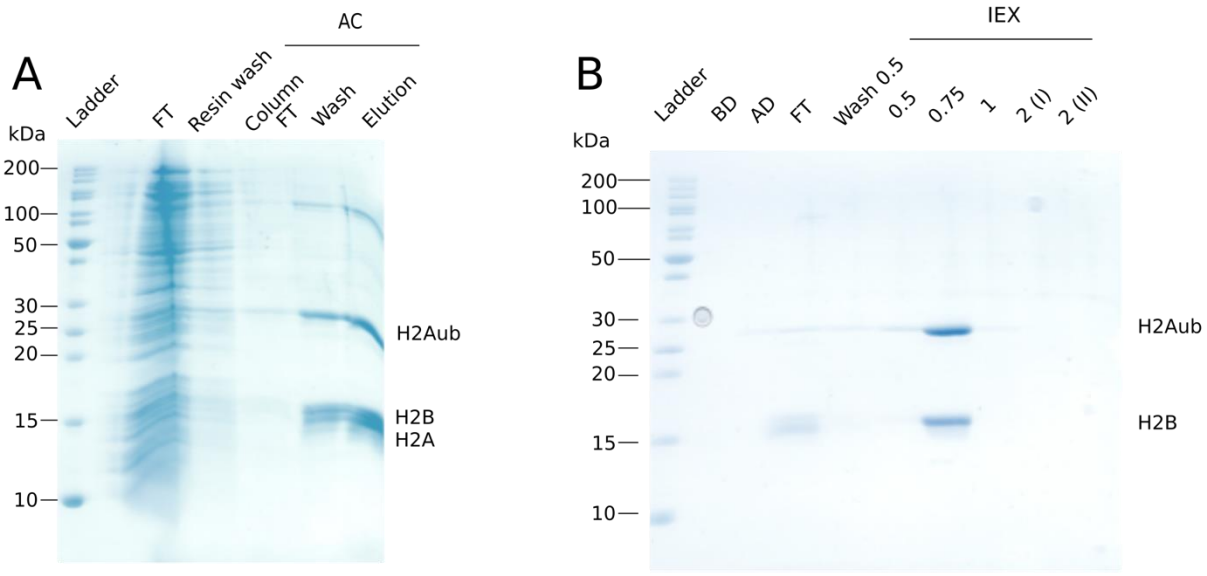


Figure 5.2 SimplyBlue™ Safe-stained 17% SDS-PAGE gels of "1093" (see Table 7.1) purification via affinity and ion-exchange chromatographies. (A) Affinity chromatography. (B) Ion-exchange chromatography. Ladder: PageRuler Unstained Protein Ladder (Thermo Scientific), FT: Flow-through; Resin wash: His lysis buffer wash; Column FT: Unbound protein after loading into the column; Wash: Washing step of the affinity chromatography; Elution: Elution step of the affinity chromatography; AC: Affinity chromatography; IEX: Ion-exchange chromatography; BD: Before dialysis; AF: After dialysis; Wash 0.5: Resin wash with 500 mM NaCl buffer; 0.5: 500 mM NaCl buffer; 0.75: 750 mM NaCl buffer; 1: 1 M NaCl buffer; 2 (I) and 2 (II): 2 M NaCl buffer (elution with 2 M NaCl buffer is executed twice). Core histones H2Aub (ubiquitinated H2A) and H2B are marked accordingly.

Table 5.1 List of purified histones.
Sample names, histones, source plasmid, expression system, and species, referenced accordingly.

Sample:	Protein:	Plasmid:	Expressed in:	Species:
XLgH3+XLgH4	WT H3 and H4 (globular)	In house	BL21 (DE3) Rosetta	*X. Laevis*
XLH3+XLH4	WT H3 and H4	In house	BL21 (DE3) Rosetta	*X. Laevis*
41	H4_I67C	In house	BL21 (DE3) Rosetta	*X. Laevis*
1102	H4_V43C and H3_F104C	In house	BL21 (DE3) Rosetta	*X. Laevis*
409B+H3	H4_A16C and WT H3	In house	BL21 (DE3) Rosetta	*X. Laevis*
1099	H4_I67C and H3_A76C	In house	BL21 (DE3) Rosetta	*X. Laevis*
1101	H4_V22C	In house	BL21 (DE3) Rosetta	*X. Laevis*
1108	H4_T67C, H3_C76A and H3_A110C	In house	BL21 (DE3) Rosetta	*X. Laevis*
1093	WT ubiquitinated H2A, and WT H2B	In house	BL21 (DE3) Rosetta	*X. Laevis*
1116	H3_L82C and H4_V81C	In house	BL21 (DE3) Rosetta	*X. Laevis*
1126	H3_L82C, H3_C110A and H4_V81C	In house	BL21 (DE3) Rosetta	*X. Laevis*
1022	Ubiquitinated H2A_ E61, H2A_D90A, H2A_E92A, and WT H2B, H3 and H4	In house	BL21 (DE3) Rosetta	*X. Laevis*

5.2. Full-length X. laevis histone octamer

To refold the histone octamer, equimolar mixing of each histone core is required. For wild-type histone octamer refolding, three different ratios of (H2A-H2B):$(H3\text{-}H4)_2$ were tested: (9:1), (10:1) and (11:1). The mixed histones were then briefly mixed and stored overnight at 4 °C. The following day, histone octamer yielding was assessed by direct visualization on a 17% SDS-PAGE gel, like depicted in Figure 5.3. Noteworthily, H2A and H2B from *X. laevis* (13.96 kDa and 13.77 kDa, respectively) are known to comigrate on SDS-PAGE gels and usually appear as a single stronger band[55]. Histone octamer assemblies using the mutant histones prepared in the current project were performed in a similar fashion. It is fulcral to know the histone:DNA ratio. If there is more histone octamer than DNA, then the excess of histones will aggregate with DNA. On the other hand, if there is more DNA than histones, then the generation of nucleosomes will be limited. Thus, titrating the histone:DNA mix is a good approach and, as suggested by Lusser and Kadonaga, varying each titration reaction by 10% is useful (e.g. histone:DNA ratios of 9:10, 10:10 and 11:10)[43,62].

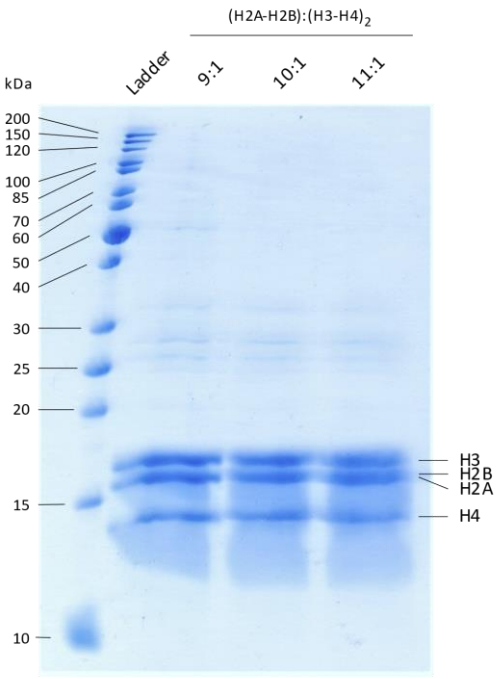


Figure 5.3 SimplyBlue™ Safe-stained 17% SDS-PAGE gel of the histone octamer with different ratios of (H2A-H2B):$(H3\text{-}H4)_2$. L: PageRuler Unstained Protein Ladder (Thermo Scientific). Three different (v/v) ratios of (H2A-H2B):$(H3\text{-}H4)_2$ are marked accordingly. Core histones (H2A, H2B, H3, and H4) are pointed out.

5.3. Synthesis and purification of nucleosomal DNA

5.3.1. Amplification by PCR

Obtaining high amounts of the long and short Widom 601 sequences which will be used in the nucleosome assembly downstream was done by PCR amplification, following the procedure described in subchapter 4.4 using the DNA template #298 (see Figure 9.2). For the extended version of the Widom 601 sequence (227 bp; +37 bp on the 5'- terminal and +44 bp on the -3' terminal), primers 797F and 797R were used. For the shorter version (149 bp) of the Widom 601 sequence, primers 423F and 423R were used (see Table 5.2 for the DNA sequences of the shorter and extended versions of the Widom 601 sequences). The size of the PCR products was checked on 1% agarose gels (see Figure 5.4).

Table 5.2 List of the two versions of the Widom 601 sequence used for nucleosome assembly.

Widom 601 version	Length (bp)	Primers	Amplified sequence via PCR (5'→ 3')
Shorter	149	423F + 423R	GCACAGGATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAATCCCCTTGGCGGTTAA AACGCGGGGGACAGCGCGTACGTGCGTTTAAGCGGTGCTAGAGCTGTCTACGACCAATTGAG CGGCCTCGGCACCGGGATTCTCCAG
Longer	227	797F + 797R	TATCCGACTGGCACCGGCAAGGTCGCTGTTCAATACATGCACAGGATGTATATATCTGACAC GTGCCTGGAGACTAGGGAGTAATCCCCTTGGCGGTTAAAACGCGGGGGACAGCGCGTACGTG CGTTTAAGCGGTGCTAGAGCTGTCTACGACCAATTGAGCGGCCTCGGCACCGGGATTCTCCA GGGCGGCCGCGTATAGGGTCCATCACATAAGGGATGAACTC

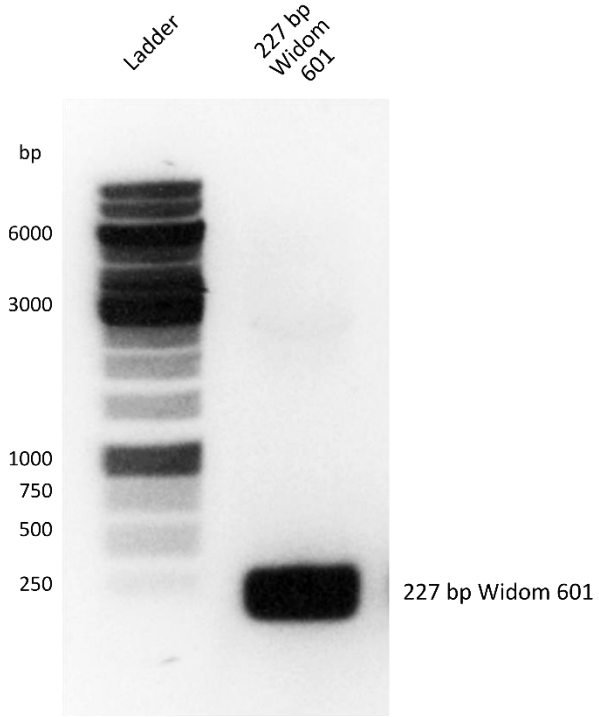


Figure 5.4 1% agarose gel of the purified 227 bp Widom 601 sequence after amplification via PCR.
Ladder: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

5.3.2. Purification by ethanol precipitation

After obtaining a considerable amount of Widom 601 sequence, purification by ethanol precipitation was ensued according to the protocol described beforehand in subchapter 4.5. Both the longer and shorter versions of the Widom 601 sequence were successfully purified (see Figure 5.5 for the longer Widom 601 sequence ethanol precipitation).

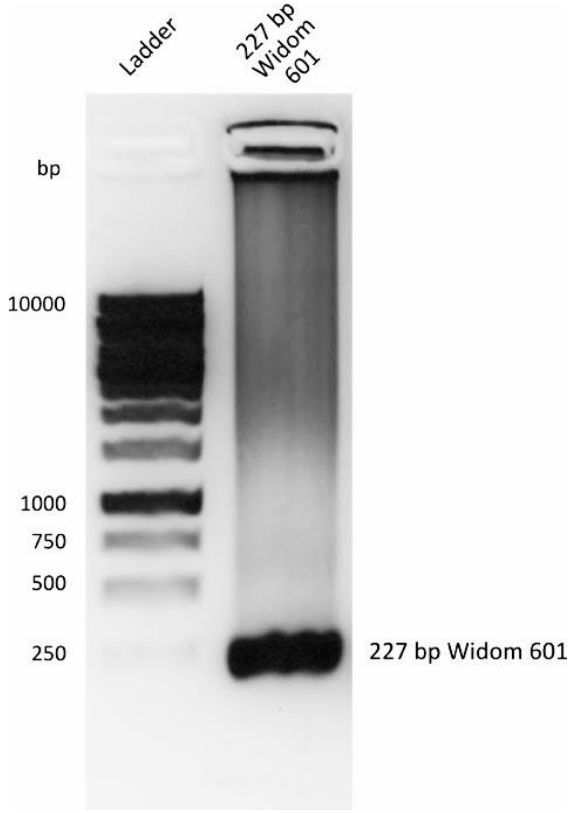


Figure 5.5 1% agarose gel of the purified 227 bp Widom 601 sequence after ethanol precipitation.
Ladder: GeneRuler 1 kb DNA Ladder (Thermo Scientific).

5.4. WT *X. laevis* nucleosome assembly

As stated in subchapter 4.9, nucleosome assembly using wild-type and mutants core histones from *X. laevis*, and the strong positioning Widom 601 sequence[63] was performed following an adapted experimental protocol based on the framework described in refs. 54–56.

As seen below in Figure 5.6, successful nucleosome assembly using WT *X. laevis* core histones and the Widom 601 sequence yields a considerable amount of centred nucleosomes, i.e. when nucleosomal DNA is properly wrapped around the histone octamer. It is not unusual to get some end-positioned nucleosomes, i.e. when the histone octamer is positioned on the edge of the nucleosomal DNA[64]. End-positioned nucleosomes migrate further than centred nucleosomes on native-PAGE electrophoresis. The thick band on top in the "Nucleosome Assembly" well is due to the presence of aggregates[65–67]. Reconstituted NCPs exhibit a characteristic slower migration in native-PAGE gels in regard to free DNA, allowing for the observation of the nucleosome:free DNA ratio[43]. As described before in subchapter 5.2, it's imperative to know the histone:DNA ratio for successful nucleosome reconstitution.

Downstream, purified nucleosomes can be used for binding and shifting assays. In this project, assembled nucleosomes were used for screening putative interactions with multiple chromatin-binding proteins via native gel shift assays and for thermal shift assays.

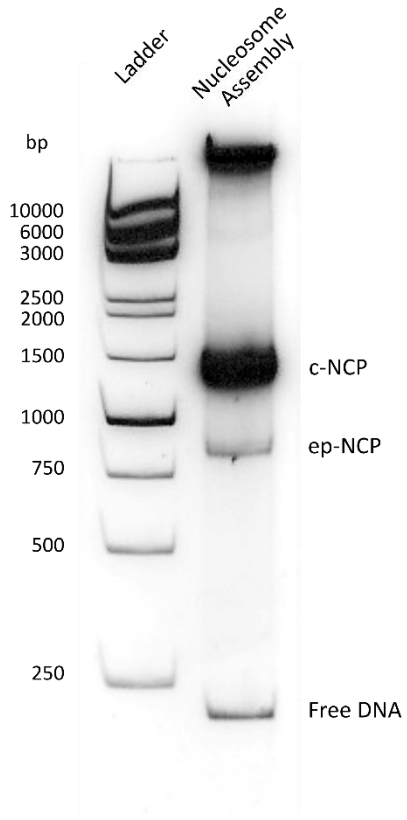


Figure 5.6 SYBR Gold-stained 6% Native-PAGE gel of the nucleosome assembly assay using WT *X. laevis* core histones and the Widom 601 sequence. Ladder: GeneRuler 1 kb DNA Ladder (Thermo Scientific); c-NCP: centred nucleosome core particle; ep-NCP: end-positioned nucleosome core particle; Free DNA: free Widom 601 sequence.

5.5. Thermal shifts of WT and mutant *X. laevis* nucleosomes from *X. Laevis*

Thermal shift assays were performed at 62 and 60 °C using WT and mutant *X. laevis* nucleosomes (for the mutant, see Figure 5.8), respectively. For each sample in the experiment, a volume of 0.5 µL of concentrated nucleosome (15 mM Tris pH 7.5, 1 mM DTT) was used. Nucleosome incubation was done in a FlexCycler^2 (Analytik Jena) PCR thermocycler. Samples were stored on ice at all times, except during incubation. Incubation times were set at 5, 15, 30 and 45 minutes. As a control, one sample was not incubated, therefore being kept on ice. As shown below in

Figure 5.7 (t = 0 min), nucleosomes tend to shift towards the DNA ends when incubated at high temperatures. Which confirms previous studies of the effect of temperature on nucleosome positioning[42,68–70]. In theory, as reported in ref. 42, this shifting would be gradually visible with time. It is plausible that, in this experiment, the temperature was high enough to allow for the swift shifting of the nucleosomes toward the DNA ends. Bands that appear lower than the ep-NCP reference band might be due to the presence of other nucleosomal states such as, for example, hexamers. The higher amount of dimer DNA might be due to nucleosome disassembly with time leading to an increase in free DNA which can then hypothetically form DNA dimers.

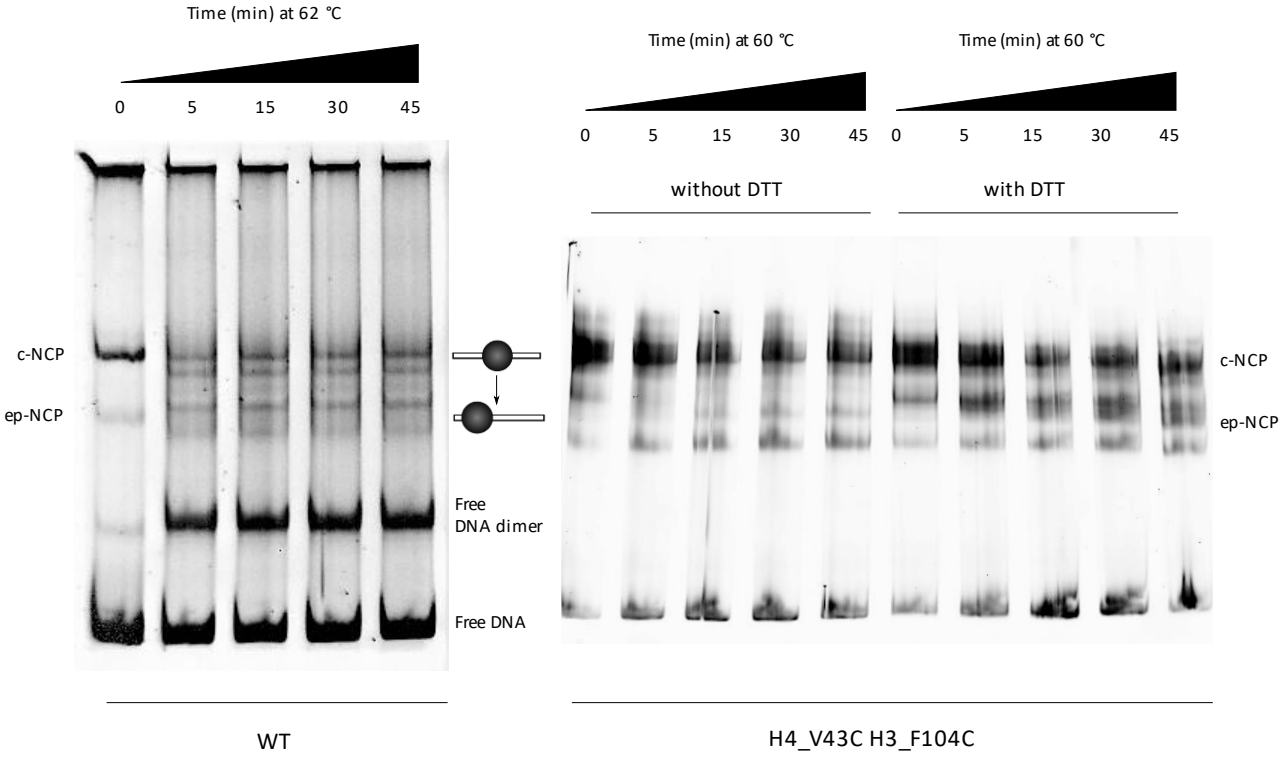


Figure 5.7 Thermal shift assays of WT and H4_V43C H3_F104C mutant *X. laevis* nucleosomes. Performed at 62 and 60 °C, respectively**.** c-NCP: centred nucleosome core particle; ep-NCP: end-positioned nucleosome core particle; Free DNA: free Widom 601 sequence; Free DNA dimer: free Widom 601 sequence dimer. Incubation times are marked accordingly in minutes.

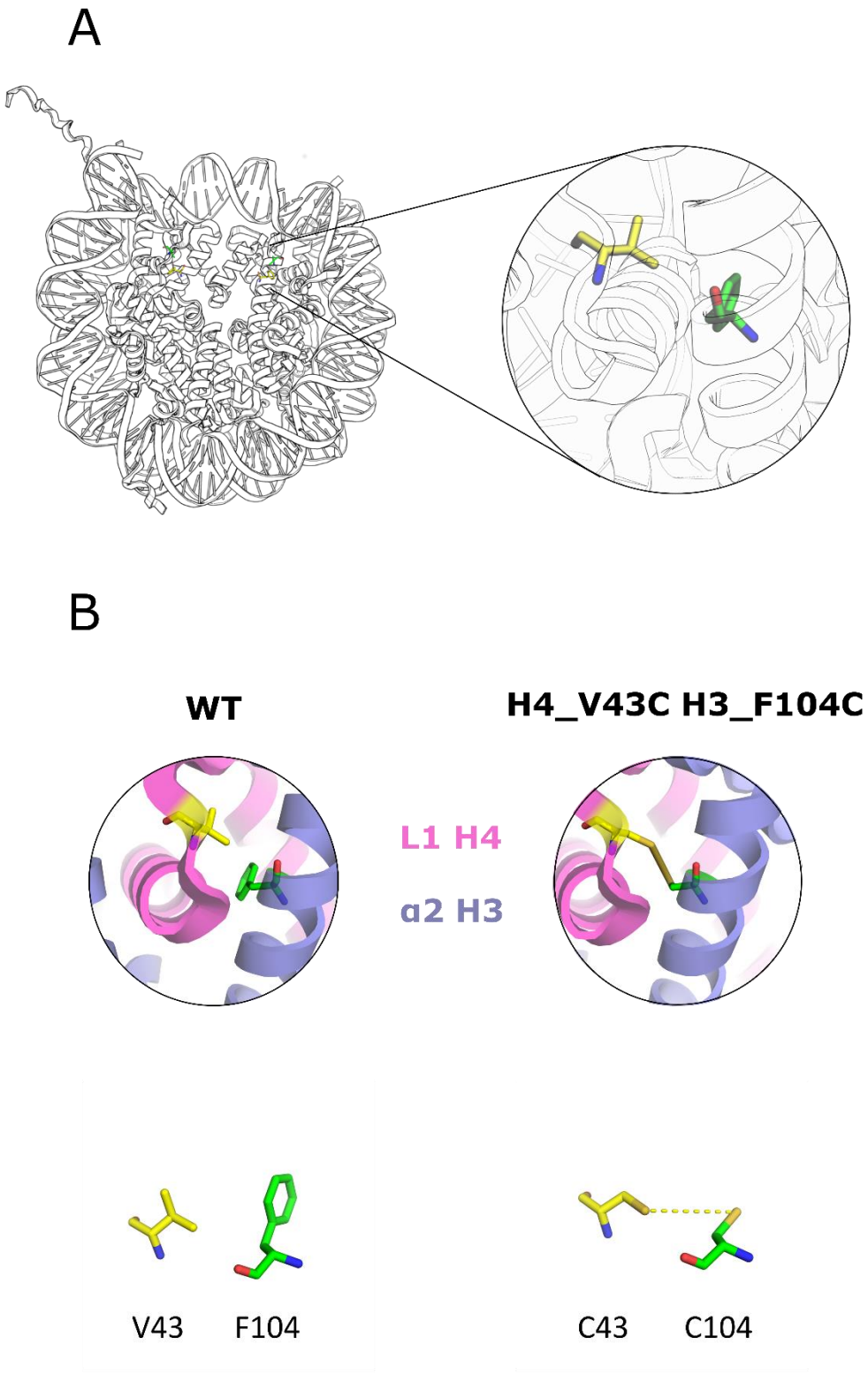


Figure 5.8 Designed *X. laevis* mutant nucleosome exhibiting H4_V43C H3_F104C mutations. (A) The nucleosome exhibiting H4_V43 and H3_F104 residues in green and yellow, respectively. (B) H4_V43C and H3_F104C mutations showing the potential to form a disulphide linkage. Representations generated using PyMOL (PDB: 1AOI; The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

5.5.1. Effect of crosslinking WT *X. laevis* nucleosomes with glutaraldehyde

Chemical crosslinking using glutaraldehyde (O=CH-CH_2-CH_2-CH_2=O), a potent extensively reported protein-protein and DNA-protein crosslinker agent (1 DNA-protein crosslink per 2.7 kb of DNA at 8 μM glutaraldehyde[71]) fixed the stable nucleosome states[72]. Crosslinking was done using 0.1% (v/v) glutaraldehyde. The control was prepared in parallel with the exact same conditions, only omitting the addition of glutaraldehyde. A SYBR Gold-stained 6% native-PAGE gel referring to the crosslinking assay is shown in Figure 5.9. As expected, the addition of 0.1% (v/v) glutaraldehyde successfully crosslinked nucleosomes since no intermediate nucleosome states are observable. The crosslinked nucleosome acts as a robust biological entity and is apparently unable to go through conformational changes. Conversely to the thermal shift assay without the addition of glutaraldehyde, the crosslinked nucleosome is unable to move and stays on the long Widom 601 sequence. This suggests that structural changes in the histone octamer move DNA.

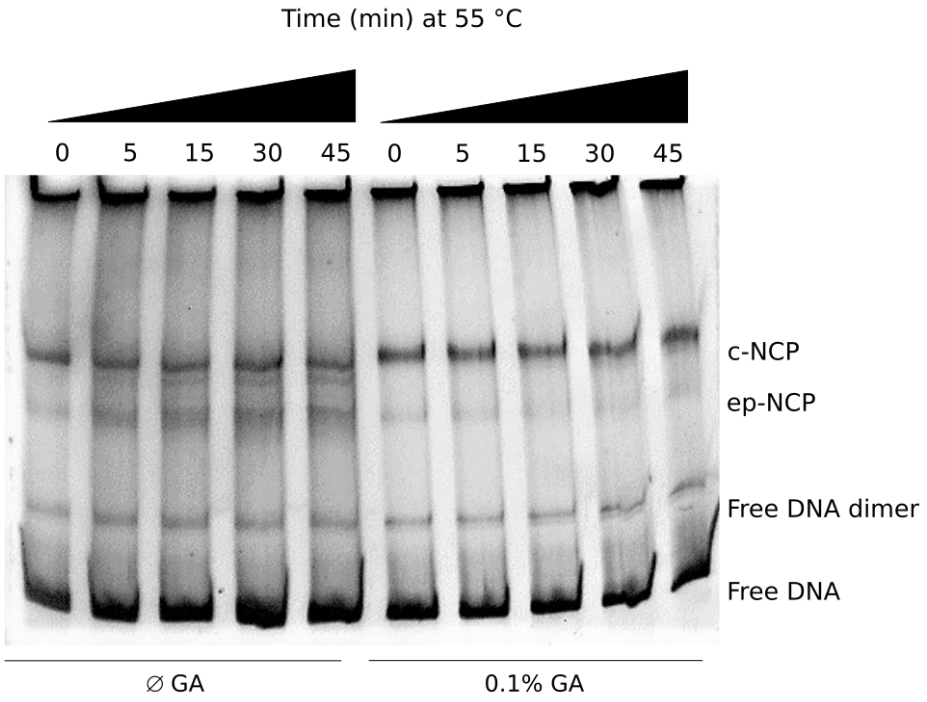


Figure 5.9 Crosslinking assay of WT *X. laevis* nucleosomes with 0.1% (v/v) glutaraldehyde at 55 °C. The control was prepared in parallel with the exact same experimental conditions but omitting the addition of glutaraldehyde. c-NCP: centred nucleosome core particle; ep-NCP: end-positioned nucleosome core particle; Free DNA: free Widom 601 sequence; Free DNA dimer: free Widom 601 sequence dimer; GA: Glutaraldehyde. Incubation times are marked accordingly in minutes.

5.6. Binding of *S. pombe* Fkbp39 to *X. laevis* core histones

Following the experiment workflow as stated in subchapter 4.11, *S. pombe* Fkbp39's ability to form a complex with WT (H2A-H2B) and $(H3\text{-}H4)_2$ from *X. laevis* was assessed. Fkpb39-SUMO-6xHis was successfully purified through affinity chromatography (6xHis-tag protein purification using Ni-NTA Agarose), although the samples contained some contaminants and Fkbp39 has been reported to copurify with some kinetochore components[61], as observed below in Figure 5.10, this was of no major concern since it will be further purified downstream.

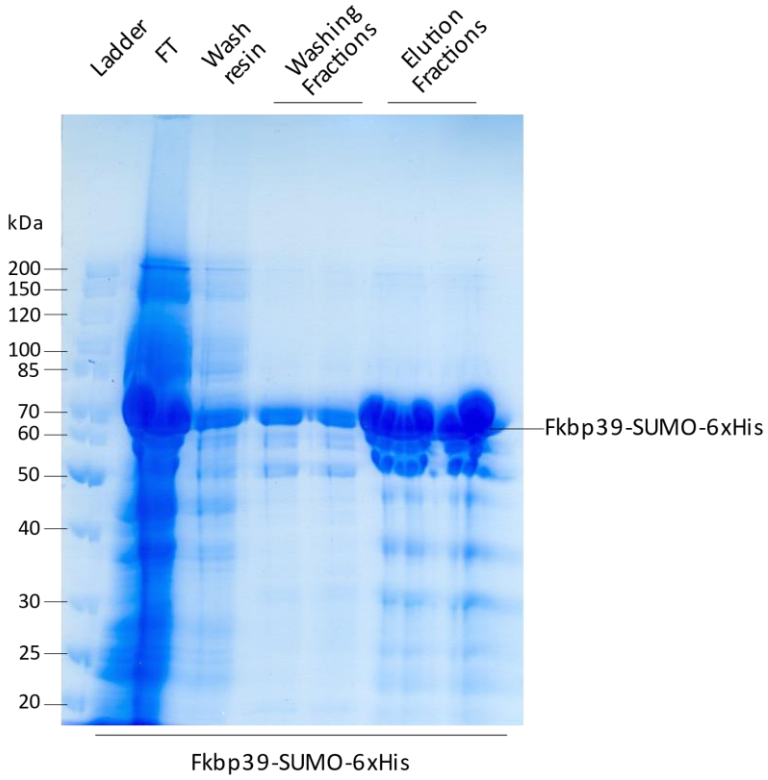


Figure 5.10 SimplyBlue™ Safe-stained 17% SDS-PAGE gel of Fkbp39-SUMO-6xHis purification via affinity chromatography. Ladder: PageRuler Unstained Protein Ladder (Thermo Scientific), FT: Flow-through, Wash resin: Resin wash with lysis wash buffer, Washing fractions: Fractions collected after adding the His washing buffer to the column; Elution fractions: Fractions collected after adding the His elution buffer to the column. Fkbp39-SUMO-6xHis tag is marked accordingly.

After confirming the enrichment and purity extent of Fkpb39-SUMO-6xHis in the elution fractions, part of it was mixed with $(H3\text{-}H4)_2$ from *X. laevis* to check for the optimal ratio of complex formation, as observed below in Figure 5.11. The resulting sample was dialysed and concentrated, after adding 10% glycerol to avoid precipitation, as depicted in Figure 5.11 (A).

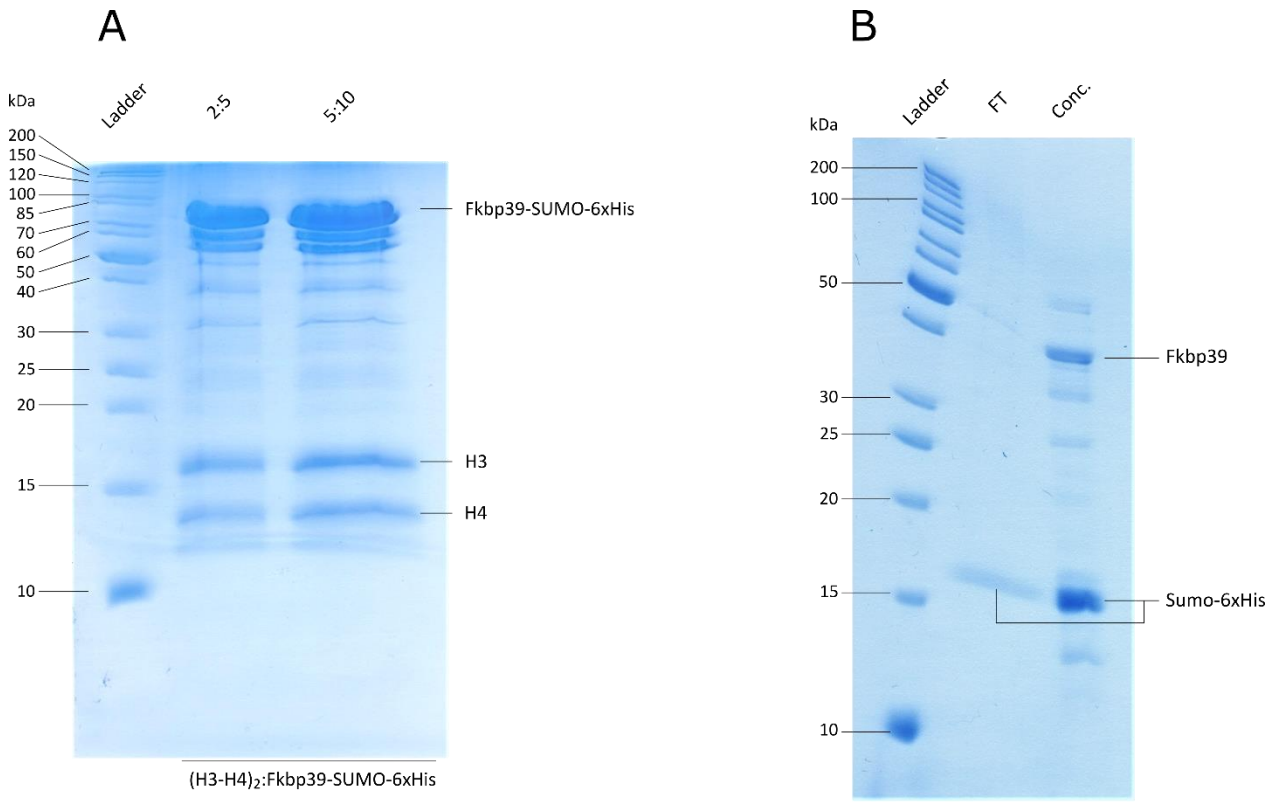


Figure 5.11 Coomassie-stained 17% SDS-PAGE gels of: (A) Fkbp39-SUMO-6xHis mixing with WT *X. laevis* $(H3\text{-}H4)_2$; and (B) Removal of the SUMO-6xHis tag from Fkbp39-SUMO-6xHis. Ladder: PageRuler Unstained Protein Ladder (Thermo Scientific), 2:5 and 5:10 are two different ratios of Fkbp39-SUMO-6xHis:$(H3\text{-}H4)_2$, FT: Flow-through after concentrating the sample, Conc.: Concentrated sample. Fkbp39-SUMO-6xHis, Fkbp39, SUMO-6xHis tag, and *X. laevis* H3 and H4 are marked accordingly.

The addition of PreScission enzyme to Fkbp39-SUMO-6xHis alone successfully allowed the removal of the SUMO-6xHis-tag, as seen above in Figure 5.11 (B). After adding the aforementioned enzyme, the sample was subdivided in three equal fractions which were then dialysed in three dialysis buffers with varying salt concentration (150, 300, and 500 mM NaCl) to determine the optimal conditions to avoid protein precipitation. From dialyses containing 150, and 300 mM NaCl, dialysed samples were pulled and further purified trough ion exchange chromatography on an ÄKTA Protein Purification System. Fractions "A6" to "B1" were analysed on a 17% SDS-PAGE gel and the most enriched fractions (A8 and A9) were mixed in stochiometric amounts with WT $(H3\text{-}H4)_2$ from the host laboratory and run on a 17% SDS-PAGE gel as well, seen below in Figure 5.12 (A) and (B), respectively.

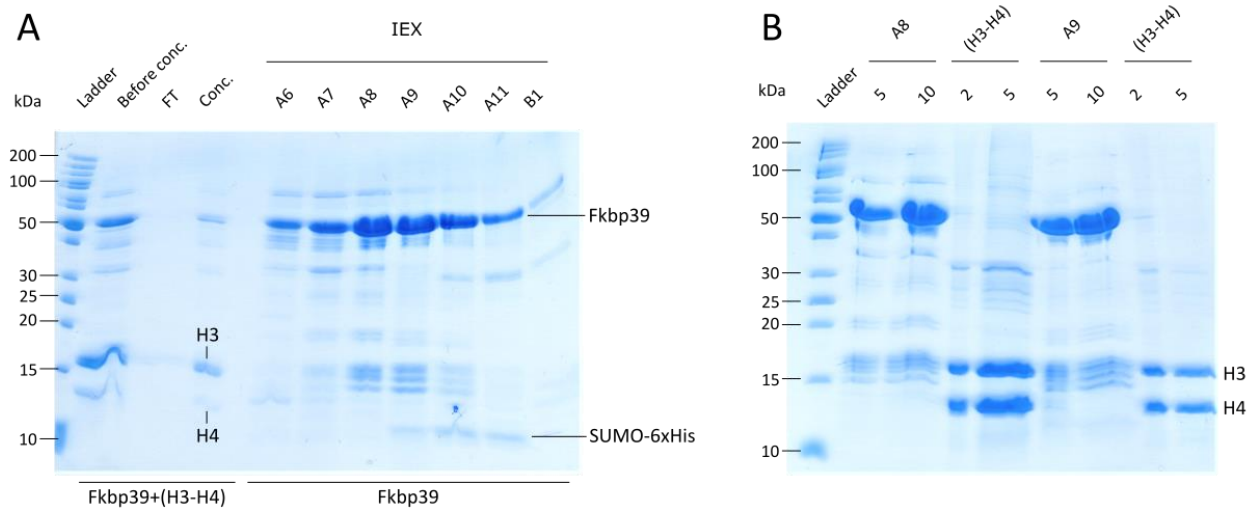


Figure 5.12 SimplyBlue™ Safe-stained 17% SDS-PAGE gels of: (A) Concentration of Fkbp39+(H3-H4)$_2$ from *X. laevis*, and Fkbp39 enriched fractions from IEX; (B) Checking optimal ratio of Fkbp39:(H3-H4)$_2$ using A8 and A9 fractions of IEX from (A) and *X. laevis* (H3-H4)$_2$. Ladder: PageRuler Unstained Protein Ladder (Thermo Scientific), Before conc.: Before concentrating the sample, FT: Concentration step flow-through, Conc.: Concentrated sample, IEX: Ion-exchange chromatography. A6, A7, A8, A9, A10, A11 and B1 are IEX collection fractions. Fkbp39, the SUMO-6xHis tag, and the *X. laevis* H3 and H4 are marked accordingly. Loaded sample volume for each sample, in µL, is noted above the respective well in (B).

After Fkbp39 and (H3-H4) were appropriately mixed (i.e. in stoichiometric amounts), 7 M guanidine hydrochloride was added and the ensuing solution was incubated and gently mixed at room temperature for 1h to allow the refolding, followed by centrifugation at 15k rpm for 15 min. The solution was then dialyzed in buffers with varying salt (NaCl) concentration and glycerol content. Dialysis in Protein dialysis buffer 200 was not successful since a considerable amount of protein precipitated (see Figure 5.13 A). Dialysis in buffers Protein dialysis buffer 200+1G, Protein dialysis buffer 200+5G, Protein dialysis buffer 200+10G, Protein dialysis buffer 300, and Protein dialysis buffer 500 yielded a good amount of complex (see Figure 5.13 B), avoiding any further losses. Noteworthily, one should avoid having glycerol in the sample if EM analyses are to be executed downstream (generally, avoid more than 1% glycerol in the sample). Samples which were dialyzed in buffers Protein dialysis buffer 300 and Protein dialysis buffer 200+1G were sent for cryo-EM analysis, and the remaining samples were frozen in liquid nitrogen and stored at -80 °C.

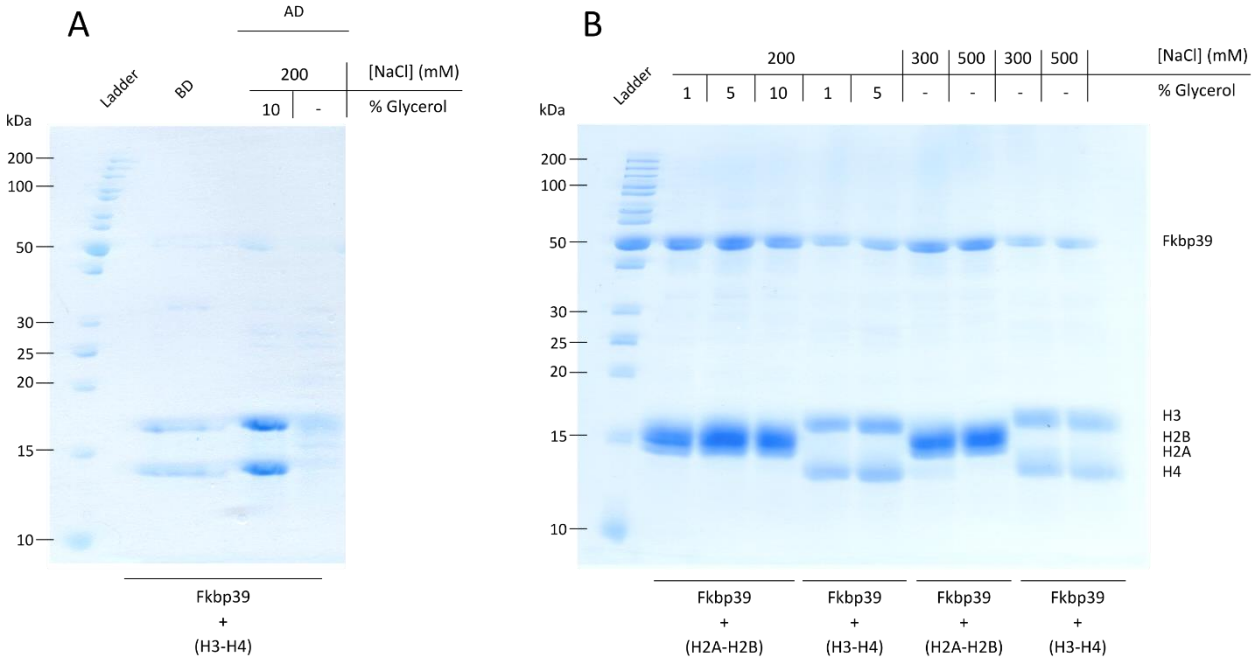


Figure 5.13 SimplyBlue™ Safe-stained 17% SDS-PAGE gels of Fkbp39+(H3-H4)$_2$ and Fkbp39+(H2A-H2B): (A) Before and after dialysis in 200 mM NaCl buffers (Protein dialysis buffer 200+10G and Protein dialysis buffer 200) of Fkbp39+(H3-H4)$_2$ from *X. laevis*; (B) Dialyzed Fkbp39+(H2A-H2B) and Fkbp39+(H3-H4)$_2$ from *X. laevis* in different dialysis buffers. Ladder: PageRuler Unstained Protein Ladder (Thermo Scientific), BD: Before dialysis; AD: After dialysis. Fkbp39 and the *X. laevis* core histones (H2A, H2B, H3 and H4) are marked accordingly. Dialysis buffer salt (NaCl) molarity (in mM) and glycerol content (in %) are marked accordingly.

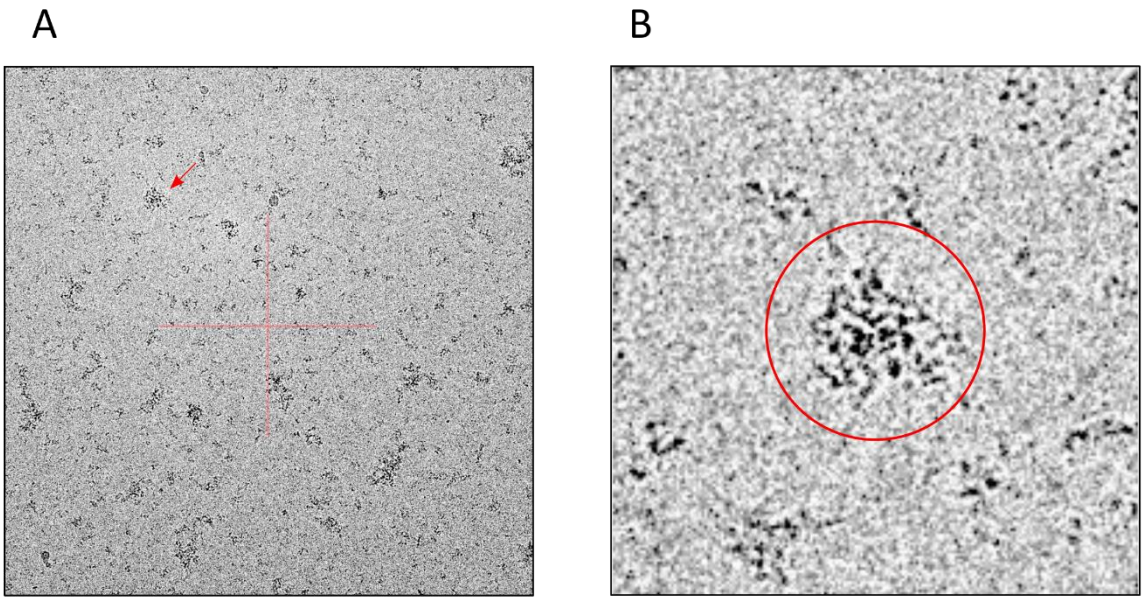


Figure 5.14 Cryo-EM snapshot of Fkbp39:(H3-H4)$_2$. (A) Aggregates seen in EM snapshot. An example is pointed by the red arrow. (B) Zoom over the pointed area in (A). Aggregates are clearly present in the sample.

Cryo-EM analyses showed that aggregation occurred, as observed in Figure 5.14. Further purifications are likely to be needed to avoid aggregate formation and allow the isolation and observation of the putative complexes. The area marked by the red arrow in Figure 5.14 A is zoomed in and shown in Figure 5.14 B to clearly show aggregate presence in the sample. Complementary cryo-EM snapshots showing the aggregates are depicted in Figure 5.15 for Fkbp39 alone, and complexed with either (H2A-H2B) or with $(H3\text{-}H4)_2$.

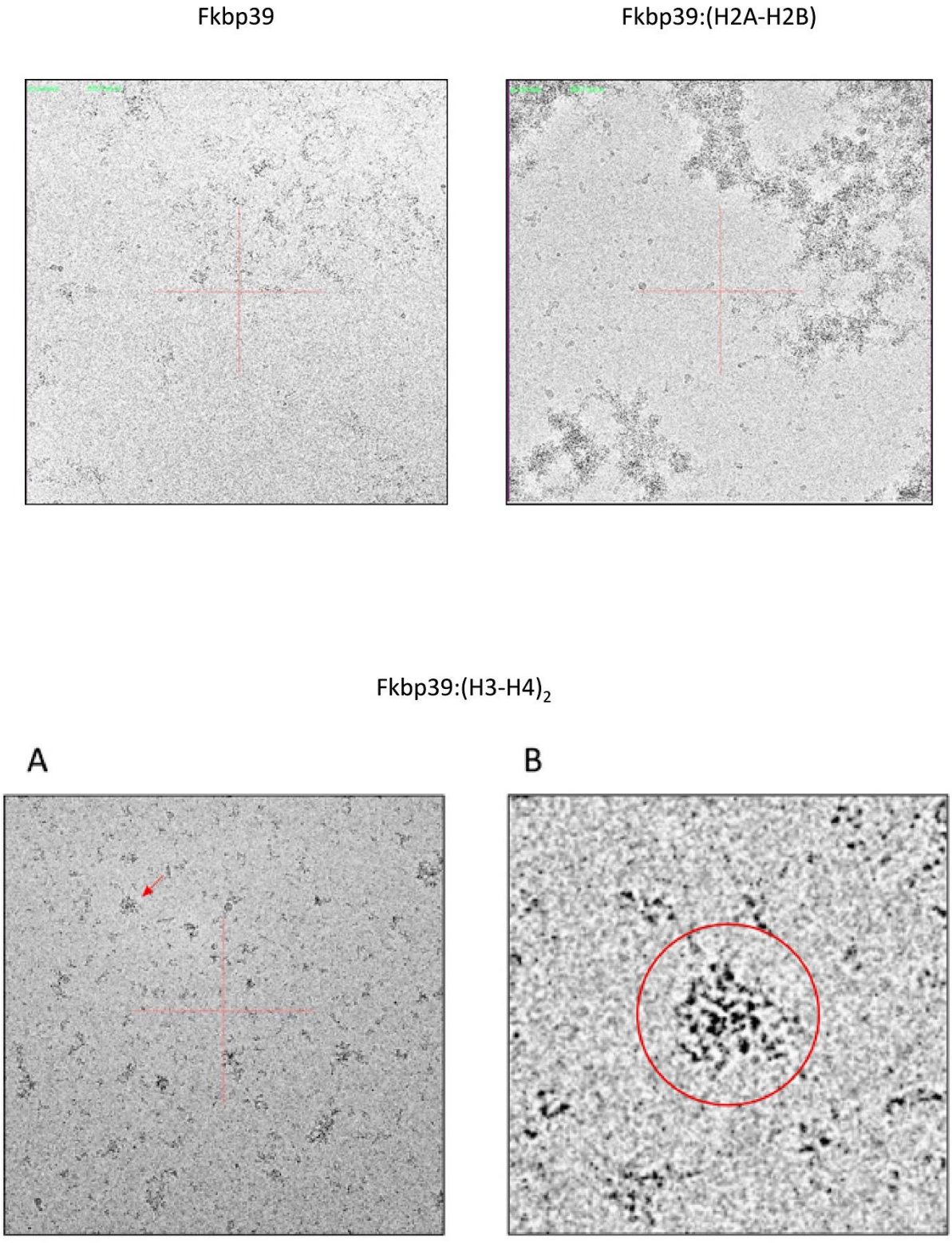


Figure 5.15 Cryo-EM snapshots of Fkbp39, Fkbp39:(H2A-H2B) and Fkbp39:(H3-H4)2.
Aggregates can be seen in all snapshots. (B) is a small area in (A) zoomed in, as marked by the red arrow. A large aggregate can be observed within the red circle in (B).

5.7. Binding affinity of chromatin-binding proteins to WT *X. laevis* nucleosomes

Although alternative experimental protocols exist, like the one described in ref. 73 which takes advantage of chromatin immunoprecipitation using H3-specific antibodies, we decided to take a simpler approach to observe putative interactions between WT *X. laevis* nucleosomes containing the Widom 601 sequence as nucleosomal DNA and chromatin-binding proteins provided by the host laboratory simply by resorting to gel mobility shift assays. The chromatin-binding proteins screened in this project are listed in Table 5.3.

Table 5.3 Chromatin-binding proteins checked for putative interactions with WT *X. laevis* nucleosomes.

Chromatin-binding protein	Gene	Full protein name	Species	Function(s)	References
Alc1	CHDL1	Chromodomain-helicase-DNA-binding protein 1-like (alternative name: amplified in liver cancer protein 1)	*H. sapiens*	• Catalyses nucleosome sliding in an ATP-dependent manner; • Nucleotide binding; • Helicase activity is strongly stimulated upon poly(ADP-ribose)-binding.	74
Alf	GTF2A1L	TFIIA-alpha and beta-like factor	*H. sapiens*	• Involved in DNA binding; • Transcription coactivator activity.	75
Hpf1	HPF1	Histone PARylation factor 1	*H. sapiens*	• Interacts with Parp1 and Parp2; • Able to change amino acid specificity toward serine; • Promotes histone serine ADP-ribosylation; • Involved in histone binding.	76,77
Hpf1	HPF1	Histone PARylation factor 1	*D. melanogaster*		
Parp1	PARP1	Poly [ADP-ribose] polymerase 1	*H. sapiens*	• Mediates the poly(ADP-ribosyl)ation of histones in a HPF1-dependent manner; • Histone deacetylase binding.	77
Parp1	PARP1	Poly [ADP-ribose] polymerase 1	*D. melanogaster*		
Parp2	PARP2	Poly [ADP-ribose] polymerase 2	*H. sapiens*	• Mediates serine ADP-ribosylation of target proteins following interaction with HPF1 (HPF1 conferring serine specificity).	76,78

Different conditions were tested to check for any shifts in the gel mobility shift assays by varying chromatin-binding protein amount, incubation temperature, incubation time, glycerol content, and using two different nucleosomal DNAs (short and long version of the Widom 601 sequence – 149 and 227 bp, respectively – prepared in this project). Furthermore, to exclude that the observed shifts were a consequence of binding between the chromatin-binding protein and DNA, an assay was executed by incubating chromatin-binding proteins with DNA only, excluding the nucleosome. To that intent, DNA was diluted accordingly to have the same concentration as in the nucleosome samples.

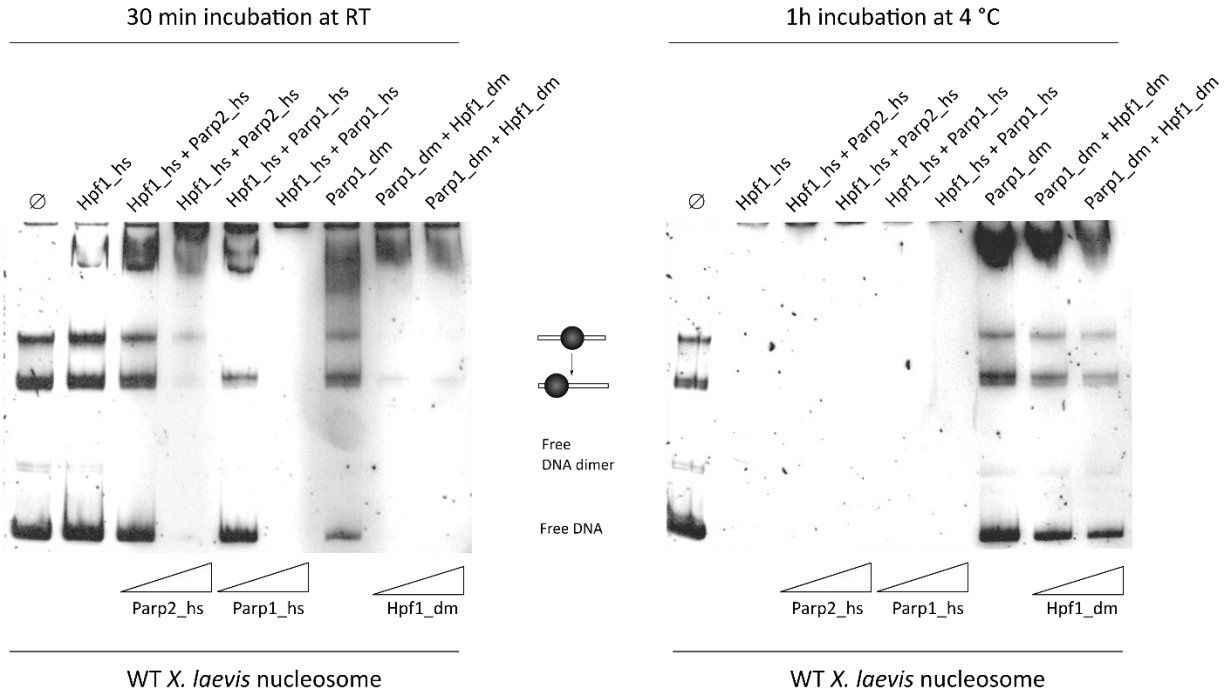


Figure 5.16 Effect of incubation time and temperature on binding efficiency of different chromatin-binding proteins.
∅: nucleosome only; hs: Homo sapiens; dm: Drosophila melanogaster.

As observed in Figure 5.16, lowering the incubation temperature, even with the increased incubation time, potentially leads to a less stable complex, since we can observe a considerable increase in aggregates. From the Figure 5.16, we can appreciate the shifts pointed out by the red arrows, indicating that *H. sapiens* Hpf1, Parp2 and Parp1, and *D. melanogaster* Parp1 might be interacting with the nucleosome. It would be interesting to see if synergetic effects exist between *H. sapiens* Hpf1 and Parp1, and Hfp1 and Parp2.

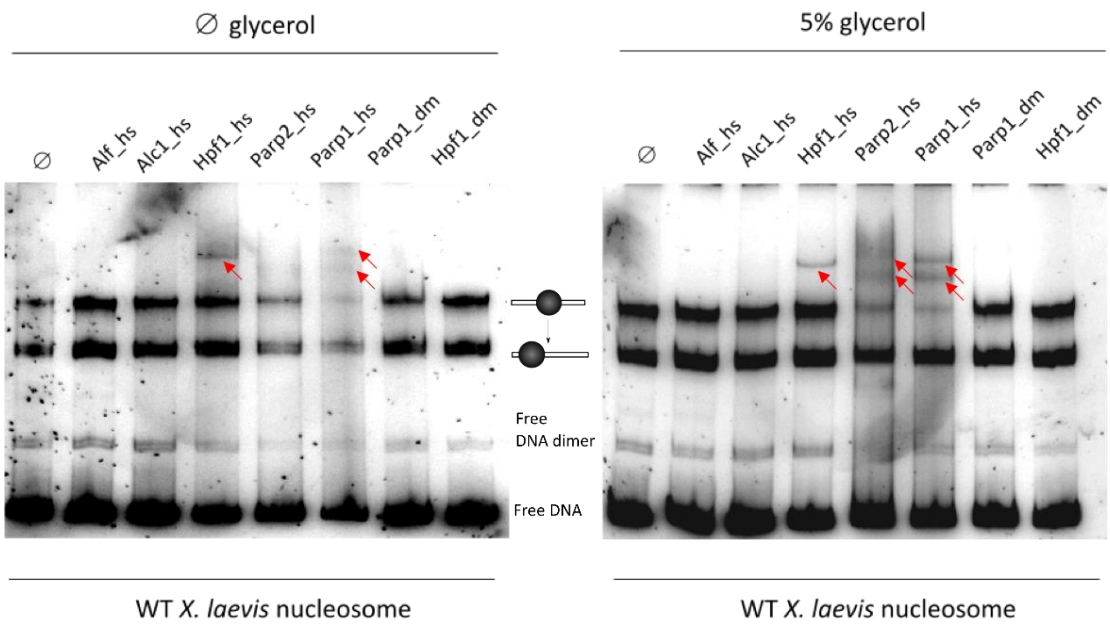


Figure 5.17 Effect of glycerol on binding efficiency of different chromatin-binding proteins.
∅: nucleosome only; hs: Homo sapiens; dm: Drosophila melanogaster.

Adding glycerol to the reaction slightly increased complex stability like depicted in Figure 5.17. Glycerol was added to avoid possible precipitation of the chromatin-binding proteins.

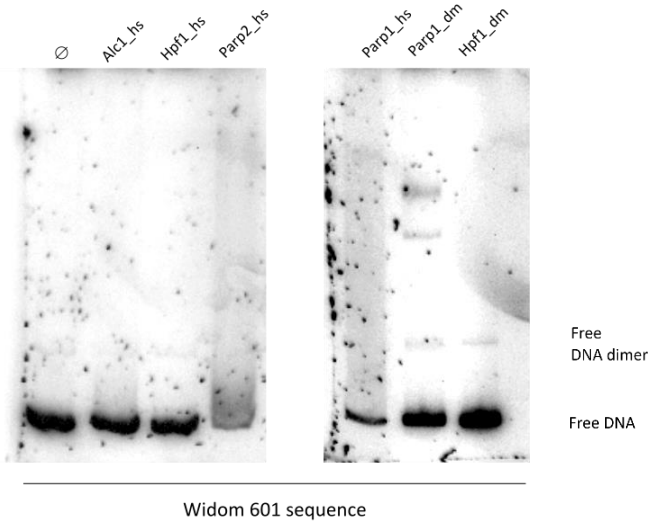


Figure 5.18 Investigating interactions between the Widom 601 DNA sequence and the chromatin-binding proteins.
∅: DNA only; hs: Homo sapiens; dm: Drosophila melanogaster.

By analysing Figure 5.18, we can assert that Parp1 and Parp2 from *H. sapiens* can hypothetically bind to the Widom 601 sequence. Moreover, *D. melanogaster* Parp1 can also recognize the aforementioned sequence. *H. sapiens* Alc1, Hpf1, and *D. melanogaster* Hpf1 do not seem to bind nucleosomal DNA alone.

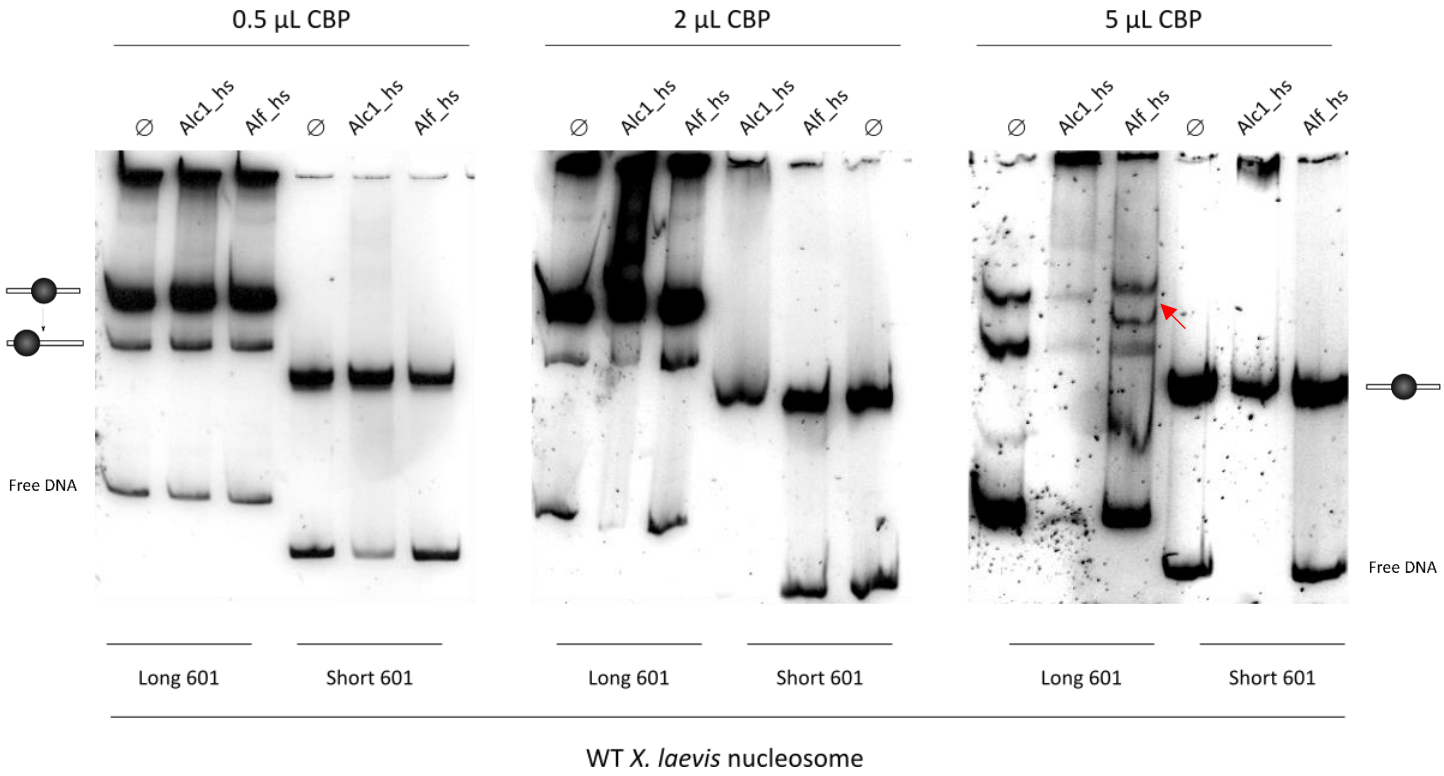


Figure 5.19 Effect of WT *X. laevis* nucleosomes with two different nucleosomal DNA sequences – long 601 (227 bp) and shorter 601 (149 bp) – on binding efficiency of *H. sapiens* Alc1 and Alf1. ∅: nucleosome only; hs: Homo sapiens; CBP: chromatin-binding protein.

A first look at Figure 5.19 leads to the assumption that shifts are only significant at higher amounts of chromatin-binding proteins, namely, at 2 µL and 5 µL. It is also apparent that *H. sapiens* Alc1 interacts with both nucleosomes, i.e. with shorter and longer 601 sequences. In regard to *H. sapiens* Alf, it seems as though it only interacts with the nucleosome which harbours the longest version of nucleosomal DNA (227 bp instead of 149 bp), as marked by the red arrow. This is a clear observation that this protein interacts with the nucleosome. It might be possible that the overhangs of the extended version of the nucleosomal DNA are required to allow proper binding of Alf to the nucleosome. Alf seems like the best candidate for further structural studies aiming at observing its binding affinity to chromatin, and possibly, understanding the binding mechanism.

Taken together, these native gel shift assays suggest that *H. sapiens* Hpf1, Parp2 and Alf are the best candidates for further structural studies aiming at defining their binding affinity the nucleosome. A good experimental technique that could be used to do so would be the ChIP-based histone association assay described in ref. 73 using H3-specific antibodies to precipitate chromatin from crosslinked whole cell extracts. Co-precipitated proteins with chromatin can be detected downstream by SDS-PAGE or Western blot analysis[73].

6. Discussion

The identification of the nucleosome, fundamental repeating subunit of chromatin, four decades ago[79] propelled chromatin research to the spotlight. With the advent of modern technologies, highly reported protocols and the availability of near-atomic resolution structures, our understanding of the nucleosomes continues to grow. However, many facets remain blurred and require further investigation to clarify the roles taken by the nucleosome in vivo. In this work, we focused on exploring the influence of histone octamer core distortion on nucleosome thermal mobility, the putative formation and cryo-EM analysis of Fkbp39:(H2A-H2B) and Fkbp39:$(H3\text{-}H4)_2$ complexes, as well as the screening of multiple chromatin-binding proteins binding affinity to the nucleosome.

6.1. Influence of histone octamer core distortion on nucleosome thermal mobility

Pioneering studies achieved in 1978 by Beard[79] using simian virus 40 minichromosomes were the first evidence of nucleosome mobility over DNA. Later, work by Flaus and Owen-Hughes [42] went further ahead and fully characterized the influence of temperature and the role of ATP on nucleosome mobility, using mouse mammary tumour virus promoter DNA as the nucleosomal DNA template. Their observations, as well as previous studies in the literature[68–70], led to the postulation that nucleosomes with short DNA fragments tend to migrate towards DNA ends during thermal-induced mobilization. This observation had obvious repercussions on gene expression, leading to the hypothesis that it would serve as an additional layer of gene expression regulation mechanism by modulating the ability of different factors to bind to chromatin.

In this work, WT and mutant nucleosome species were prepared to observe the effect of temperature on nucleosome mobility as well as the influence of the designed mutations. Therefore, several *X. laevis* histones (WT and mutant) were firstly purified after heterologously expressing them in *E. coli* BL21 (DE3) Rosetta strain. Core histones H2A and H2B, and core histones H3 and H4 were co-expressed. Co-expression of histones is very practical and reduces time-consuming individual histone purification protocols, producing soluble complexes that can be readily purified downstream[80], e.g. via affinity chromatography. With regard to nucleosomal DNA, an extended version (227 bp) of the previously reported strong-positioning Widom 601 sequence[39] was synthesized in high amounts via PCR and then purified by EtOH precipitation. Nucleosome core particles were assembled using an adapted protocol based on the ones stated in refs. 54–56 (workflow shown in Figure 1.10).

Firstly, our thermal shift assays at 62 ºC (with incubation times of 5, 15, 30 and 45 minutes) using WT nucleosomes corroborate the previously reported thermally-driven nucleosome mobility of WT species (see

Figure 5.7). We observe the same tendency to migrate towards the DNA extremities when incubated at higher temperatures. Hence, we wanted to access the effect of the designed mutations in the mutant nucleosome species over its mobility at high temperatures. Our H4_V43C H3_F104C mutant nucleosome (checked through sequencing) shows the typical mobility pattern at 60 ºC (shifting of the nucleosome towards the DNA ends, i.e. end-positioned nucleosome). In the case of our mutant, two different conditions were tested: with DTT and without DTT. As expected, shifting towards the extremities was more marked when nucleosomes were incubated in reducing conditions (with DTT) as seen by the presence of stronger bands below the "c-NCP" band, corresponding to the different end-positioned nucleosomes ("ep-NCP"). DTT, being a reducing agent, will reduce disulphide bonds, and

naturally, disulphide bridges present between two cysteine residues will be dismantled. One would then expect a higher degree of mobility when incubated at high temperatures for some time as the result of weakened intramolecular and intermolecular forces, as observed in this case. Interestingly, without DTT, although some shifting is still observable, it is clearly not as marked as in comparison with nucleosomes incubated in reducing conditions (with DTT). And it seems like longer incubation times do not affect shifting to a greater extent in both contexts. This suggests that these mutations (H4_V43C, H3_F104C) significantly affect nucleosome structural integrity, most possibly by strengthening the overall structure through the increasing of the intermolecular forces most possibly through the establishment of a disulphide bridge at the interface between L1 (loop 1) of core histone H4 and α2 (α-helix 2) of core histone H3, hence hindering its thermally-driven mobility. This suggests that histone octamer plasticity is vital for nucleosome to slide along DNA in a noncatalyzed manner.

Chemical crosslinking is a very fit experimental method for protein-protein interaction identification. In theory, the crosslinking agent leads to the formation of covalent bonds between two proteins by reacting with the functional groups of the protein. In the cellular context, many transient interactions between proteins occur. In vitro, when working with purified proteins, this technique is most useful for validating the presence of protein-protein interactions. Not only can crosslinking be used to show that two proteins are in close proximity to each other and form interactions, but it can also give insights on which regions the contacts occur[81]. Crosslinking of WT *X. laevis* nucleosomes with 0.1% (v/v) glutaraldehyde was successful and led to the formation of stable crosslinked nucleosomes with no observable intermediate nucleosome states and very minimal end-positioned (e-NCP) nucleosomes in comparison to centred nucleosomes (c-NCP, see Figure 5.9), suggesting that crosslinked histone octamers are immobilized on DNA, thus indicating that the structural changes in the histone octamer move the underlying DNA[82]. Increasing incubation time had no effect on nucleosome mobility since no variation in end-positioned or intermediate state nucleosomes was observed. Incubation temperature was set at 55 ºC since it is sufficient to observe nucleosome mobility as seen in the case of non-crosslinked nucleosomes in Figure 5.9.

The intrinsic plasticity of the histone octamer allows the uncatalyzed DNA translocation and is a prerequisite for multiple chromatin remodelling complexes. Apart from DNA translocation and nucleosome flexibility, structural rearrangements of the histone octamer likely influence the stability of the different conformational states of the nucleosome, which might in turn be used by chromatin remodelers at their advantage to execute their function(s).

6.2. Fkbp39 structural studies

FK506-binding protein 39 kDa, abbreviated “Fkbp39” (and alternatively, “Ani1”), is a PPiase with a conserved NPL domain that acts as a histone chaperone. This protein is involved in proline isomerization on core histone H3 N-terminal tail affecting H3 methylation, thus having a role in gene expression regulation[20,58].

In this work, we aimed at purifying *S. pombe* Fkbp39 for downstream cryo-EM analysis. Furthermore, we also tried to check for complex formation with (H2A-H2B) and $(H3\text{-}H4)_2$, and to observe the putative complexes under cryo-EM at the Max Planck Institute of Biochemistry (Munich, Germany).

The purification step of Fkbp39 required a SUMO tag to increase solubility and stability, as well as a 6xHis tag to allow for affinity chromatography purification. The SUMO-6xHis was successfully cut out with the addition of *in house* PreScission enzyme. Initially, purification of Fkbp39:$(H3\text{-}H4)_2$ was not successful since it only yielded significant amounts in buffers with more than 1 % glycerol, which restricted analysis by cryo-EM. Hence, different dialysis buffers were prepared with varying salt (200, 300 and 500 mM NaCl) molarities and different glycerol content (0, 1, 5 and 10 %), which resulted in high yields of the putative Fkbp39:(H2A-H2B) and Fkbp39:$(H3\text{-}H4)_2$ complexes. The best-looking samples (dialyzed in buffers Protein dialysis buffer 300 and Protein dialysis buffer 200+1G) were sent for cryo-EM analysis. Unfortunately, isolation and visualization of either Fkbp39 alone or Fkbp39 complexed with core histones (H2A-H2B) or with $(H3\text{-}H4)_2$ was unsuccessful due to aggregate formation. Further optimization strategies are likely to be required to acquire more stable assemblies. Fkbp39’s intrinsically disordered regions might be preventing the formation of stable complexes with core histones.

6.3. Chromatin-binding proteins screening

To check for binding of the multiple putative chromatin-binding proteins (listed under Table 5.3) to WT *X. laevis* nucleosomes, we resorted to native gel shift assays – a "rapid and sensitive method to detect protein-nucleic acid interactions"[83].

Prior to running the electrophoreses, samples containing WT *X. laevis* nucleosomes mixed with the chromatin-binding protein(s) were incubated at room temperature for 30 min, or at 4 ºC for 1h. Furthermore, glycerol content was also varied. Both these settings were used to check for the optimal reaction conditions to allow proper visualization of shifts, if occurring. Increasing incubation time from 30 min to 1h and reducing incubation temperature from room temperature to 4 ºC mostly led to aggregate formation (see Figure 5.16). Adding 5% glycerol, on the other hand, led to stronger shift bands (marked in by red arrows in Figure 5.17), possibly by decreasing the formation of aggregates.

Looking at Figure 5.17, it is clear that Hpf1 (*H. sapiens*), Parp1 (*H. sapiens*) and Parp2 (*H. sapiens*) interact with the nucleosome. To exclude the possibility of interaction of these proteins with the Widom 601 sequence alone (as free nucleosomal DNA), an additional native gel shift assay was done incubating the tested proteins with DNA alone (see Figure 5.18). Interestingly, Parp1 (*H. sapiens*) and Parp2 (*H. sapiens*) seem to bind to DNA alone, as well as Parp1 (*D. melanogaster*). Further experiments are required to fully disclose the binding affinities to either free nucleosomal DNA, or the nucleosome itself.

A first glance at Figure 5.16 provides a snapshot of the binding affinity of the different tested proteins to the nucleosome. Different combinations of proteins were tested and screened for binding with the nucleosome in this assay: one chromatin-binding protein alone or a combination of proteins with varying amounts. As indicated by the red arrows, Hpf1 (*H. sapiens*) and Parp1 (*D. melanogaster*) look like good candidates for further structural studies aiming at their affinity to chromatin. Additionally, combining Hpf1 (*H. sapiens*) with increasing amounts of Parp2 or Parp1 (*H. sapiens*) further increases shifting, though this experiment alone does not provide enough information, i.e. the interactions between the aforementioned proteins would have to be checked first, and then synergetic effects might also be worth considering.

Taking a closer look at Alc1 (*H. sapiens*) and Alf (*H. sapiens*) in Figure 5.19 (gel with 5 µL of chromatin-binding protein, on the far right), using native gel shift assays with the short (149 bp nucleosomal DNA) and long (227 bp nucleosomal DNA) nucleosomes, Alc1 quite strikingly interacts with both the long and short nucleosomes and Alf interacts with the long nucleosome only.

Further assays will be required to unambiguously prove the binding of the best chromatin-binding candidates observed in this project (i.e. *H. sapiens* Hpf1, Parp2 and Alf), notwithstanding that these experiments served as a good means to get a glimpse at which protein might be the best contender to study its interactions with the nucleosome.

7. Conclusion

The main aim of this project was to observe the influence of histone octamer core distortion on nucleosome thermal mobility. Here, we corroborate the shifting of WT *X. laevis* nucleosomes towards the extremities (end-positioned) in regards to the nucleosomal DNA (using a 227 bp extended version of the strong positioning Widom 601 sequence) when incubated at high temperatures (62 ºC)[42]. Moreover, this typical shifting pattern of nucleosomes is also reflected in the double mutant nucleosome (H4_V43C H3_F104C) assembled in this work (incubated at 60 ºC). Interestingly, adding reducing conditions (via the addition of DTT) when incubating the double mutant yielded a high amount of intermediate nucleosomal states, conversely to the assay without DTT. Taken together, these observations suggest that these mutations (H4_V43C and H3_F104C) strengthen the overall nucleosome structure at the interface between L1 (loop 1) of core histone H4 and α2 (α-helix 2) of core histone H3, hence affecting its thermally-driven mobility by limiting nucleosome sliding along nucleosomal DNA and suggesting that histone octamer plasticity is a prerequisite for nucleosome to move along DNA in a noncatalyzed fashion.

Notwithstanding that purification of Fkbp39 (*S. Pombe*) was successful, Fkbp39:(H2A-H2B) and Fkbp39:$(H3\text{-}H4)_2$ putative complexes analyses by cryo-EM failed due to aggregate formation. Further protocol optimization strategies are likely to improve complex formation conditions and might stabilize the putative complexes, thus allowing for downstream structural analysis.

Lastly, the screening of the putative chromatin-binding proteins binding to WT *X. laevis* nucleosomes suggested that Alc1, Alf, Hpf1, Parp1 and Parp2 (*H. sapiens*) might interact with the nucleosome. Most intriguingly, Hpf1, Parp2 and Alf (*H. sapiens*) seem like the best candidates for further studies focused on studying their ability to bind to nucleosomes. A good approach to do so would be to use the ChIP-based histone association assay described in ref. 73 which uses H3-specific antibodies to precipitate chromatin from crosslinked whole cell extracts. Crosslinking causes the co-precipitation of chromatin with the bound proteins which can then be detected by SDS-PAGE or Western blot analysis[73].

8. References

1. Kornberg, R. D. & Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* (1999). doi:10.1016/S0092-8674(00)81958-3
2. Alva, V., Ammelburg, M., Söding, J. & Lupas, A. N. On the origin of the histone fold. *BMC Struct. Biol.* (2007). doi:10.1186/1472-6807-7-17
3. Cosgrove, M. S., Boeke, J. D. & Wolberger, C. *Regulated nucleosome mobility and the histone code*. *Nature Structural and Molecular Biology* **11,** 1037–1043 (Nature Publishing Group, 2004).
4. Workman, J. L. & Abmayr, S. M. *Fundamentals of chromatin*. *Fundamentals of Chromatin* (2014). doi:10.1007/978-1-4614-8624-4
5. Li, B., Carey, M. & Workman, J. L. The Role of Chromatin during Transcription. *Cell* **128,** 707–719 (2007).
6. Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D. & Klug, A. Structure of the nucleosome core particle at 7 resolution. *Nature* (1984). doi:10.1038/311532a0
7. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* (1997). doi:10.1038/38444
8. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* (2002). doi:10.1016/S0022-2836(02)00386-8
9. Bartke, T. *et al.* Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* **143,** 470–484 (2010).
10. Annunziato, A. DNA Packaging: Nucleosomes and Chromatin. *Nat. Educ.* **1,** 1 (2008).
11. Ramakrishnan, V. Histone structure and the organization of the nucleosome. *Annu. Rev. Biophys. Biomol. Struct.* **26,** 83–112 (1997).
12. Edlich-Muth, C. *et al.* The pentameric nucleoplasmin fold is present in Drosophila FKBP39 and a large number of chromatin-related proteins. *J. Mol. Biol.* **427,** 1949–1963 (2015).
13. Gurard-Levin, Z. A., Quivy, J.-P. & Almouzni, G. Histone Chaperones: Assisting Histone Traffic and Nucleosome Dynamics. *Annu. Rev. Biochem.* **83,** 487–517 (2014).
14. PomBase - Gene - ani1 (SPBC1347.02) - CENP-A amino terminus domain (NTD) isomerase Ani1. Available at: https://www.pombase.org/gene/SPBC1347.02. (Accessed: 20th March 2018)
15. Wood, V. *et al.* PomBase: A comprehensive online resource for fission yeast. *Nucleic Acids Res.* **40,** (2012).
16. Matsuyama, A. *et al.* ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. *Nat. Biotechnol.* **24,** 841–7 (2006).
17. Koztowska, M. *et al.* Nucleoplasmin-like domain of FKBP39 from Drosophila melanogaster forms a tetramer with partly disordered tentacle-like C-terminal segments. *Sci. Rep.* **7,** 1–14 (2017).
18. Pfam: Protein: FKBP4_SCHPO (O74191). Available at: http://pfam.xfam.org/protein/O74191. (Accessed: 20th March 2018)
19. FPR3 / YML074C Overview. Available at: https://www.yeastgenome.org/locus/S000004539. (Accessed: 20th March 2018)
20. O74191 (FKBP4_SCHPO). Available at: http://www.uniprot.org/uniprot/O74191. (Accessed: 20th March 2018)
21. Bateman, A. *et al.* UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45,** D158–D169 (2017).
22. Park, S.-K., Xiao, H. & Lei, M. Nuclear FKBPs, Fpr3 and Fpr4 affect genome-wide genes transcription. *Mol. Genet. Genomics* **289,** 125–136 (2014).
23. Ramos, I. *et al.* The intrinsically disordered distal face of nucleoplasmin recognizes distinct oligomerization states of histones. *Nucleic Acids Res.* **42,** 1311–25 (2014).
24. Hondele, M. & Ladurner, A. G. The chaperone-histone partnership: For the greater good of histone traffic and chromatin plasticity. *Curr. Opin. Struct. Biol.* **21,** 698–708 (2011).

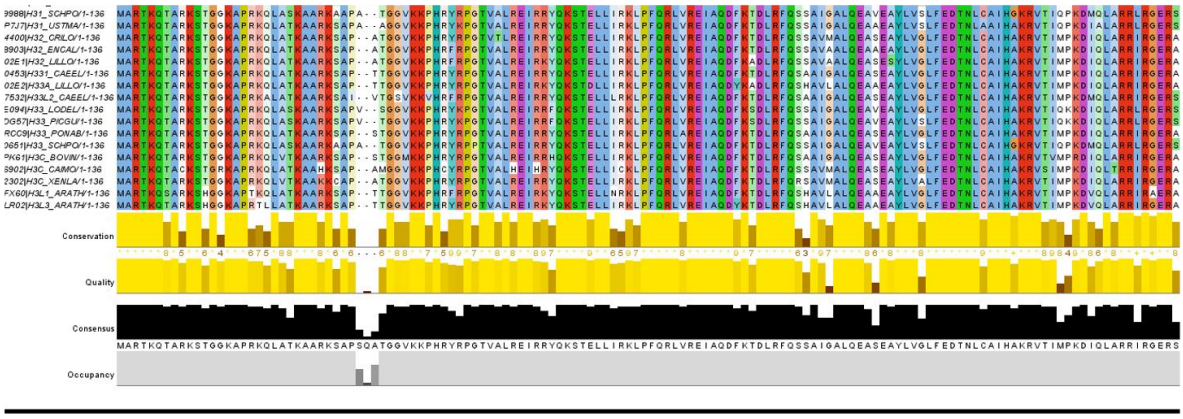
25. Luger, K. Structure and dynamic behavior of nucleosomes. *Curr. Opin. Genet. Dev.* **13,** 127–135 (2003).
26. Cosgrove, M. S. & Wolberger, C. How does the histone code work? *Biochem. Cell Biol.* **83,** 468–476 (2005).
27. De La Cruz, X., Lois, S., Sánchez-Molina, S. & Martínez-Balbás, M. A. Do protein motifs read the histone code? *BioEssays* **27,** 164–175 (2005).
28. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128,** 693–705 (2007).
29. Bowman, G. D. & Poirier, M. G. Post-translational modifications of histones that influence nucleosome dynamics. *Chemical Reviews* **115,** 2274–2295 (2015).
30. Zentner, G. E. & Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nature Structural and Molecular Biology* **20,** 259–266 (2013).
31. Thomas Jenuwein, C. D. A. Translating the Histone Code. (2001).
32. Henikoff, S. & Shilatifard, A. Histone modification: Cause or cog? *Trends in Genetics* **27,** 389–396 (2011).
33. Cui, L. & Miao, J. Chromatin-Mediated epigenetic regulation in the malaria parasite Plasmodium falciparum. *Eukaryotic Cell* **9,** 1138–1149 (2010).
34. Lee, J., Smith, E. & Shilatifard, A. The Language of Histone Crosstalk Jung-Shin. **142,** 682–685 (2013).
35. Sinha, K. K., Gross, J. D. & Narlikar, G. J. Distortion of histone octamer core promotes nucleosome mobilization by a chromatin remodeler. *Science (80-.).* (2017). doi:10.1126/science.aaa3761
36. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119,** 941–953 (2004).
37. Sun, Z. W. & Allis, C. D. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418,** 104–108 (2002).
38. Bowman, G. D. Mechanisms of ATP-dependent nucleosome sliding. *Curr. Opin. Struct. Biol.* **20,** 73–81 (2010).
39. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276,** 19–42 (1998).
40. Frouws, T. D., Duda, S. C. & Richmond, T. J. X-ray structure of the MMTV-A nucleosome core. *Proc. Natl. Acad. Sci.* **113,** 1214–1219 (2016).
41. Andrews, A. J. & Luger, K. Nucleosome Structure(s) and Stability: Variations on a Theme. *Annu. Rev. Biophys.* **40,** 99–117 (2011).
42. Flaus, A. & Owen-Hughes, T. Dynamic Properties of Nucleosomes during Thermal and ATP-Driven Mobilization. *Mol. Cell. Cell Biol.* **23,** 7767–7779 (2003).
43. Lusser, A. & Kadonaga, J. T. Strategies for the reconstitution of chromatin. *Nat. Methods* **1,** 19–26 (2004).
44. Ito, T., Tyler, J. K. & Kadonaga, J. T. Chromatin assembly factors: a dual function in nucleosome formation and mobilization? *Genes Cells* **2,** 593–600 (1997).
45. Tatchell, K. & Van Holde, K. E. Nucleosome Reconstitution: Effect of DNA Length on Nucleosome Structure. *Biochemistry* **18,** 2871–2880 (1979).
46. Ruiz-Carrillo, A. & Jorcano, J. L. An Octamer of Core Histones in Solution: Central Role of the H3·H4 Tetramer in the Self-Assembly1. *Biochemistry* **18,** 760–768 (1979).
47. Kornberg, R. D. & Thonmas, J. O. Chromatin Structure: Oligomers of the Histones. *Science (80-.).* **184,** 865–868 (1974).
48. Simon, R. H. & Felsenfeld, G. A new procedure for purifying histone pairs H2A+H2B and H3+H4 from chromatin using hydroxylapatite. *Nucleic Acids Res.* **6,** 689–696 (1979).
49. Khuong, M. T., Fei, J., Cruz-Becerra, G. & Kadonaga, J. T. A simple and versatile system for the ATP-dependent assembly of chromatin. *J. Biol. Chem.* **292,** 19478–19490 (2017).
50. Rhodes, D. & Laskey, R. A. Assembly of nucleosomes and chromatin in Vitro. *Methods Enzymol.* **170,** 575–585 (1989).
51. Huynh, V. A. T., Robinson, P. J. J. & Rhodes, D. A method for the in vitro reconstitution of a defined '30 nm' chromatin fibre containing stoichiometric amounts of the linker histone. *J. Mol. Biol.* **345,** 957–968 (2005).
52. Tanaka, Y. *et al.* Expression and purification of recombinant human histones. *Methods* **33,** 3–

11 (2004).
53. Luger, K., Rechsteiner, T. J. & Richmond, T. J. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.* **304,** 3–19 (1999).
54. Ivić, N., Groschup, B., Bilokapić, S. & Halić, M. Simplified Method for Rapid Purification of Soluble Histones. *Croat. Chem. Acta* **89,** 1–10 (2016).
55. Rogge, R. A. *et al.* Assembly of Nucleosomal Arrays from Recombinant Core Histones and Nucleosome Positioning DNA. *J. Vis. Exp.* 1–10 (2013). doi:10.3791/50354
56. Luger. Methods for Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA. **375,** 23–44 (2004).
57. Klinker, H., Haas, C., Harrer, N., Becker, P. B. & Mueller-Planitz, F. Rapid purification of recombinant histones. *PLoS One* (2014). doi:10.1371/journal.pone.0104029
58. Kuzuhara, T. & Horikoshi, M. A nuclear FK506-binding protein is a histone chaperone regulating rDNA silencing. *Nat. Struct. Mol. Biol.* (2004). doi:10.1038/nsmb733
59. Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P. & Ranson, N. A. An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods* **100,** 3–15 (2016).
60. 4X SDS-PAGE loading buffer. *Cold Spring Harb. Protoc.* **2006,** pdb.rec10588-pdb.rec10588 (2006).
61. Przewloka, M. R. *et al.* Molecular analysis of core kinetochore composition and assembly in Drosophila melanogaster. *PLoS One* **2,** (2007).
62. Fyodorov, D. V. & Kadonaga, J. T. Dynamics of ATP-dependent chromatin assembly by ACF. *Nature* **418,** 896–900 (2002).
63. Vasudevan, D., Chua, E. Y. D. & Davey, C. A. Crystal Structures of Nucleosome Core Particles Containing the '601' Strong Positioning Sequence. *J. Mol. Biol.* **403,** 1–10 (2010).
64. Isaac, R. S. *et al.* Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. *Elife* **5,** (2016).
65. Bertin, A., Mangenot, S., Renouard, M., Durand, D. & Livolant, F. Structure and phase diagram of nucleosome core particles aggregated by multivalent cations. *Biophys. J.* **93,** 3652–3663 (2007).
66. De Frutos, M., Raspaud, E., Leforestier, A. & Livolant, F. Aggregation of nucleosomes by divalent cations. *Biophys. J.* **81,** 1127–1132 (2001).
67. Allahverdi, A., Chen, Q., Korolev, N. & Nordenskiöld, L. Chromatin compaction under mixed salt conditions: Opposite effects of sodium and potassium ions on nucleosome array folding. *Sci. Rep.* **5,** (2015).
68. Meersseman, G., Pennings, S. & Bradbury, E. M. Mobile nucleosomes - a general behavior. *EMBO J* **11,** 2951–2959 (1992).
69. Flaus, A. & Richmond, T. J. Positioning and stability of nucleosomes on MMTV 3'LTR sequences. *J. Mol. Biol.* **275,** 427–441 (1998).
70. Sakaue, T., Yoshikawa, K., Yoshimura, S. H. & Takeyasu, K. Histone core slips along DNA and prefers positioning at the chain end. *Phys. Rev. Lett.* **87,** 78105-1-78105–4 (2001).
71. Kuykendall, J. R. & Bogdanffy, M. S. Efficiency of DNA-histone crosslinking induced by saturated and unsaturated aldehydes in vitro. *Mutat. Res. Lett.* **283,** 131–136 (1992).
72. Czarnota, G. J. & Ottensmeyer, F. P. Structural states of the nucleosome. *J. Biol. Chem.* **271,** 3677–3683 (1996).
73. Ricke, R. M. & Bielinsky, A. K. Easy detection of chromatin binding proteins by the histone association assay. *Biol. Proced. Online* **7,** 60–69 (2005).
74. Ahel, D. *et al.* Poly(ADP-ribose)-dependent regulation of DNA repair by the chromatin remodeling enzyme ALC1. *Science (80-.).* (2009). doi:10.1126/science.1177321
75. Upadhyaya, A. B., Lee, S. H. & Dejong, J. Identification of a general transcription factor TFIIAalpha/beta homolog selectively expressed in testis. *J. Biol. Chem.* (1999). doi:10.1074/jbc.274.25.18040
76. Bonfiglio, J. J. *et al.* Serine ADP-Ribosylation Depends on HPF1. *Mol. Cell* (2017). doi:10.1016/j.molcel.2017.01.003
77. Gibbs-Seymour, I., Fontana, P., Rack, J. G. M. & Ahel, I. HPF1/C4orf27 Is a PARP-1-Interacting Protein that Regulates PARP-1 ADP-Ribosylation Activity. *Mol. Cell* (2016).

doi:10.1016/j.molcel.2016.03.008
78. Amé, J. C. *et al.* PARP-2, a novel mammalian DNA damage-dependent poly(ADP-ribose) polymerase. *J. Biol. Chem.* (1999). doi:10.1074/jbc.274.25.17860
79. Beard, P. Mobility of histones on the chromosome of simian virus 40. *Cell* (1978). doi:10.1016/0092-8674(78)90279-9
80. Anderson, M. *et al.* Co-expression as a convenient method for the production and purification of core histones in bacteria. *Protein Expr. Purif.* (2010). doi:10.1016/j.pep.2010.03.013
81. Kapoor, M. How to crosslink proteins. *Cell. Mol. Microb. Biol. Div. …* 1–6 (1996).
82. Bilokapic, S., Strauss, M. & Halic, M. Structural rearrangements of the histone octamer translocate DNA. *Nat. Commun.* (2018). doi:10.1038/s41467-018-03677-z
83. Hellman, L. M. & Fried, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat. Protoc.* (2007). doi:10.1038/nprot.2007.249
84. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).
85. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25,** 1189–1191 (2009).

9. Appendix

A



B

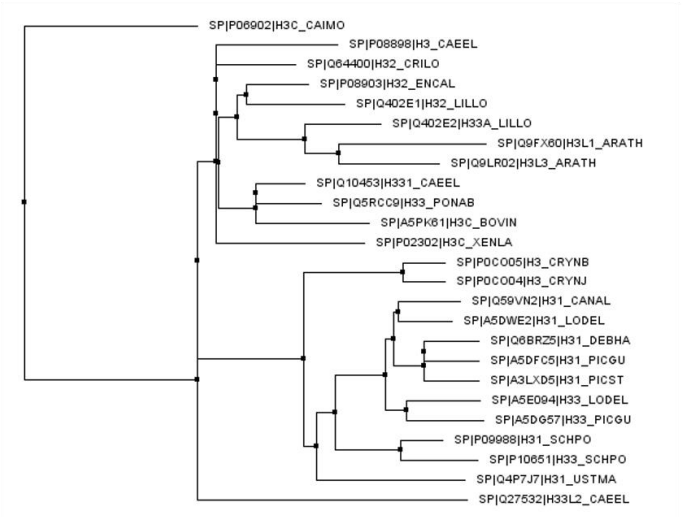


Figure 9.1 Multiple sequence alignment of a histone H3 cluster in eukaryotes. The cluster (cluster ID: UniRef90_P08898) was obtained through UniProt database, using manually refined entries only (Swiss-Prot). (A) Visualization of the alignment was performed via Jalview 2.10.3b1 using the Clustal color scheme (blue: hydrophobic, red: positive charge, magenta: negative charge, green: polar, pink: cysteines, orange: glycines, yellow: prolines, cyan: aromatic, white: unconserved). Conservation, quality, consensus and occupancy bar plots indicate the extent of: conservation, likelihood of observing mutations, percentage of the modal residue and ungapped positions for each amino acid residue (i.e. per column), respectively. (B) The phylogenetic tree was built using the neighbor joining method with the BLOSUM62 matrix on Jalview 2.10.3b1[85].

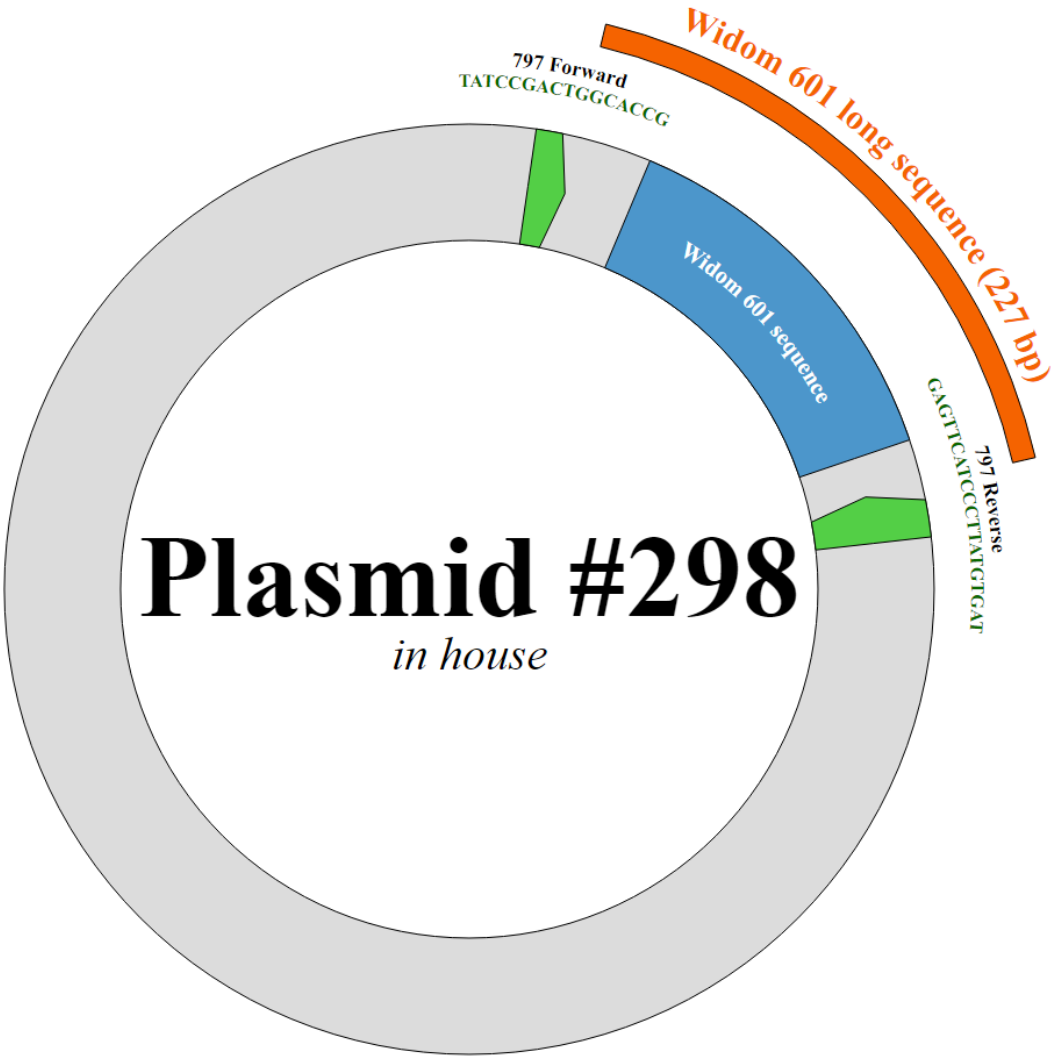


Figure 9.2 PCR amplification schematic of the long version of the strong position Widom 601 sequence (227 bp) using 797F and 797R primers with #298 plasmid template. Canonical Widom 601 sequence shown in blue. Long Widom 601 sequence (227 bp) shown in orange. *In house* plasmid sequence #298 sequence is undisclosed. Representation using AngularPlasmid (http://angularplasmid.vixis.com/).